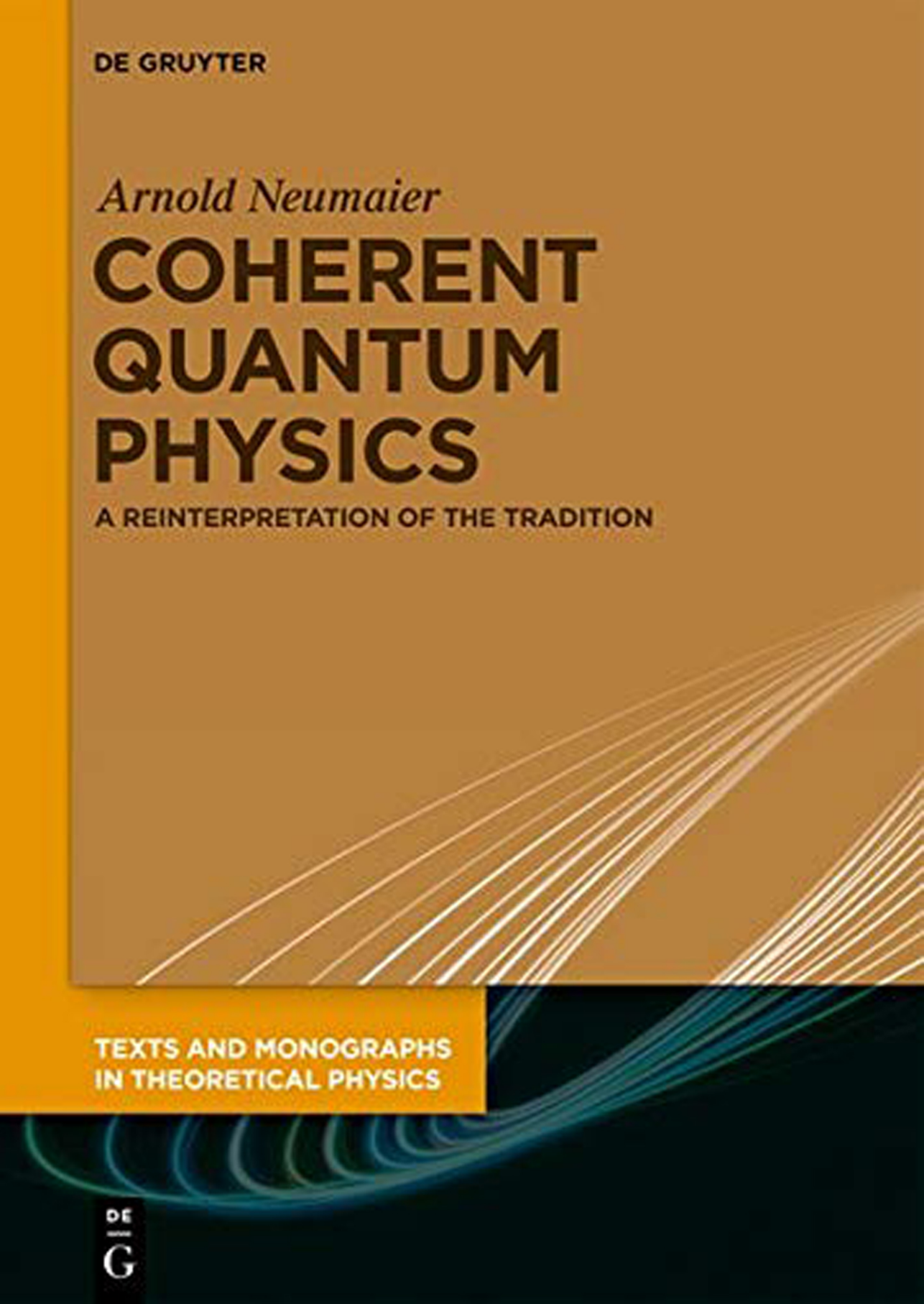
DE GRUYTER

Arnold Neumaier

COHERENT QUANTUM PHYSICS

A REINTERPRETATION OF THE TRADITION



TEXTS AND MONOGRAPHS IN THEORETICAL PHYSICS

DE G

Arnold Neumaier

Coherent Quantum Physics

Texts and Monographs in Theoretical Physics

Edited by
Michael Efroimsky, Bethesda, Maryland, USA
Leonard Gamberg, Reading, Pennsylvania, USA

Arnold Neumaier

Coherent Quantum Physics

A Reinterpretation of the Tradition

DE GRUYTER

Mathematics Subject Classification 2010
Primary: 81P15, 81R30, 46E22; Secondary: 17B81, 81T99

Author
Prof. Dr. Arnold Neumaier
Universität Wien
Fakultåt für Mathematik
Oskar-Morgenstern-Platz 1
1090 Wien
Austria
Arnold.Neumaier@univie.ac.at

ISBN 978-3-11-066729-5
e-ISBN (PDF) 978-3-11-066738-7
e-ISBN (EPUB) 978-3-11-066736-3
ISSN 2627-3934

Library of Congress Control Number: 2019947573

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de.

Cover image: Guy N Harris / iStock / Getty Images Plus
Typesetting: VTeX UAB, Lithuania
Printing and binding: CPI books GmbH, Leck

www.degruyter.com

To Maria,
in honor of the Creator of our magnificent universe

Preface

In a statistical description of nature only expectation values or correlations are observable.
Christof Wetterich, 1997 [299, p. 2678]

One is almost tempted to assert that the usual interpretation in terms of sharp eigenvalues is 'wrong', because it cannot be consistently maintained, while the interpretation in terms of expectation values is 'right', because it can be consistently maintained.
John Klauder, 1997 [160, p. 6]

What has become known as the quantum measurement problem [...] encapsulates many of the fundamental conceptual difficulties that have to this date prevented us from arriving at a commonly agreed-upon understanding of the physical meaning of the formalism of quantum mechanics and of how this formalism relates to the perceived world around us.
Maximilian Schlosshauer, 2007 [265, p. VIII]

This book introduces mathematicians, physicists, and philosophers to a new, coherent approach to theory and interpretation of quantum physics (including quantum mechanics, quantum statistical mechanics, quantum field theory, and their applications), in which classical and quantum thinking live peacefully side by side and jointly fertilize the intuition.

An interpretation of quantum mechanics relates its formalism to the actual informal practice of using quantum mechanics in our scientific culture. An impeccable interpretation must show that there is a fully consistent relation between theory and practice. The interpretation may use concepts familiar from our culture to explain the working of quantum physics in practice to everyone's satisfaction.

What are the shortcomings of the current approaches? The minimal statistical interpretation predicts the statistics of outcomes of experiments. It is silent about the interpretation of quantum mechanics in the absence of measurements, and therefore about the interpretation of quantum physics applied to the far past of the universe, before experiments were possible. This constitutes a serious gap—the interpretation is consistent, but incomplete (as it should be for a "minimal" interpretation). The Copenhagen interpretation, which claims that nothing can be asserted in the absence of a measurement, is also consistent. But this sounds like the concept that a tree fallen in the wood has fallen only after someone has seen it. This is one of the reasons why quantum mechanics comes across as somewhat strange. In a many-world interpretation, the world splits and splits, completely unnoticed by us, into all possible futures. This is science fiction by conception. The other known interpretations are either variations of the above or require additional, in principle, unobservable, and hence fictional stuff. As a result, much of quantum physics appears to the general public as a kind of quantum magic.

Why do physicists live with this? A noteworthy aspect of the standard interpretations is that the state vector cannot represent the whole universe, since it must

https://doi.org/10.1515/9783110667387-201

exclude an observer or measuring device that determines when a measurement has occurred. This is the so-called Heisenberg cut between the quantum and the classical world. To date, this has not been a problem in making successful experimental predictions, so practitioners are often satisfied with the quantum formalism in a standard interpretation. Tradition builds the quantum edifice on a time-honored foundation that accounts for essentially all experimental facts. But it takes a "shut-up-and-calculate" attitude towards the interpretation of the foundations. The traditional presentation of quantum physics is clearly adequate for prediction, but seems not to be suitable for an adequate understanding.

A second reason is that a number of popular "quantum magicians", very experienced quantum physics practitioners specializing in quantum optics, like to give their audience the impression that important parts of quantum mechanics are weird. And the general public loves it! Part of the magicians' art consists of remaining silent about the true reasons why things work rationally, since then the weirdness is gone, and with it the entertainment value.

Does quantum mechanics have to be weird? It sells much better to the general public if it is presented that way, and there is a long history of proceeding like this. But it is an obstacle for everyone who wants to truly understand quantum mechanics, and to physics students, who have to unlearn what they were told as laypersons. When presented in the right way, quantum mechanics is not at all weird, but very close to classical mechanics. Much of the weirdness comes from forcing quantum mechanics into the straightjacket of a particle picture. The particle picture breaks down completely in the subatomic domain, as witnessed by the many weird things that result from such a view.

Coherent quantum physics removes the radical split between classical mechanics and quantum mechanics. This book demonstrates that at any level of detail, Nature can be rationally and objectively understood just by interpreting the traditional, well-established mathematics of quantum physics in an appropriate way. This requires a reinterpretation of the tradition. The interpretation featured in this book succeeds without any change in the theory, and without introducing new counterintuitive features or new theoretical concepts. The resulting quantum features then are only those familiar from everyday life.

Nature, as we perceive it with our eyes, consists of images—in mathematical terms 2-dimensional fields, with properties (colors) at each point. Our brains interpret images as scenes in a, strictly speaking, not directly perceived 3-dimensional world of objects. The same object seems larger or smaller depending on its distance from us, with a shape that is deduced from images showing the object from different perspectives. All our observations are indirect: We perceive images and other sensory information and infer the true (theoretical, reproducible, invariant) properties of the objects around us.

From the experience of the multitude of such sensory perceptions of many people, our culture created a network of concepts and relations now called science, and in

particular physics. Space has become 3-dimensional, represented at each particular time by 3-dimensional fields that tell the spatial properties of the materials present at each point in space. Their boundaries delineate the objects, some sharply, others—such as clouds—only in a fuzzy way.

Space thus becomes equipped with many properties. There are local properties, such as temperature, colors, hardness, stress, and chemical composition. In fluids there are properties like salt concentration, but also pressure, streaming velocity, et cetera. Each of these gives rise to a field that specifies how these properties vary with the position in space. In addition, there are less tangible invisible properties, such as those described by the electromagnetic field. The latter describes the properties responsible for the electric and magnetic phenomena in Nature, on which much of our modern culture depends. Additionally, there are bilocal properties, such as distances between two points in space. There are also nonlocal, region-dependent properties, such as the diameter, mass, and volume of an extended object, or the surface area of its boundary.

Objects often move. Just like photographs of stars in a long term night exposure, they trace out tracks in an abstract 3-dimensional space. These tracks form curves of a thickness depending on the objects' size. The theory of special relativity teaches us beyond this 3-dimensional picture of the world a 4-dimensional perspective in a 4-dimensional Minkowski space, whose coordinates represent both space and time. Due to length contraction and time dilation, shapes look and clocks move differently for observers moving at different velocities relative to each other. In special relativity, moving points are represented by so-called world lines; the curves they trace out in Minkowski space. The objects we see are extended in space, and therefore trace out world tubes—thin or thick tracks in 4 dimensions with boundaries reflecting the sharp or fuzzy, constant or changing shape of the objects.

Materials vibrate and produce sound. The electromagnetic field vibrates and produces light. Both are phenomena characterizing the behavior of waves. These can be decomposed into harmonic waves of specified direction and frequency. The possible frequencies of vibration make up a spectrum. A small part of these spectra are directly observable by the human ear and eye; a very large part is indirectly observable through various spectroscopic techniques.

Fields are representations of the continuum, infinitely divisible space and time. But continuous fields are also the cause of discrete events. Continuous water waves may cause discrete, random damage. Bullets fired on plexiglass described by the stress fields of continuum mechanics cause visible, discrete random cracks emanating from the center of impact. Casting a die, modeled by the continuous laws of classical mechanics, results in a random, discrete value—depending on which face it falls.

If we compare the motion of the Moon, a car, a leaf falling from a tree, or a pollen corn in water, we realize that light objects move less predictably. This introduces a second form of randomness into scientific descriptions. Often, measurements do not produce exactly the same results. Typically, the best empirical approximation to the true

value of something measured is a simple average of multiple measurements—there is a democracy of measurement results. This insight, a form of the law of large numbers, justifies statistical techniques. They allow one to obtain much useful information from many inaccurate measurements.

This feature of Nature extends down to the smallest scales. On the scale of human experience, unanimated matter is highly predictable. But on the molecular and atomic level, matter is observed to behave mostly in a random way. Therefore, the reproducible information about microscopic events consists mostly of statistical properties, such as chemical reaction rates. On the subatomic level, Nature's behavior is so uncertain that even the opinions on what exists are somewhat controversial.

A new approach. From a more technical perspective, the new approach described in this book may be summarized as follows:

Coherent quantum physics is physics in terms of a coherent space consisting of a line bundle over a classical phase space and an appropriate coherent product. The kinematical structure of quantum physics and the meaning of the fundamental quantum observables are given by the symmetries of this coherent space, their infinitesimal generators, and associated operators on the quantum space of the coherent space.

The formal, mathematical core of quantum physics is cleanly separated from the interpretational issues. To achieve this, we need to avoid some of the traditional quantum mechanical jargon. In particular, following the convention of Allahverdyan et al. [7], we add the prefix "q-" to all traditional quantum notions that suggest by their name a particular interpretation, and hence might confuse the borderline between theory and interpretation. In particular, the operators usually called[1] "observables" will be called "q-observables" to distinguish them from observables in the operational sense of numbers obtainable from observation. Similarly, we use the terms q-expectation and q-probability for the conventional but formally defined terms expectation and probability.

Objective properties, including their uncertainties are given by q-expectations of products of quantum fields and what is computable from these. The dynamics of the universe is given by the Ehrenfest equations for q-expectations, and defines the dynamics of every physical subsystem by restriction. Particles are approximate effective descriptions of certain extended blops of mass and/or energy, descriptions that make sense only under special conditions.

Certain q-expectations are approximately observed in experiments. Like ordinary averages, q-expectations become more accurate (that is, less uncertain) by averaging over many similar items. Averaging over macroscopic spacetime regions produces macroscopic quantities with negligible uncertainty, and leads to classical physics.

---

1 This notion appears first in Dirac's 1930 book [70, pp. 28]. Later editions make the restriction that observables are Hermitian, and have real spectrum.

The new approach involves one radical step, the reinterpretation of an assumption underlying traditional quantum physics that was virtually never questioned before: The eigenvalue link between theory and observation is replaced by a q-expectation link. This leads to a new interpretation of quantum mechanics, the **thermal interpretation**, introduced in Chapter 9. It transforms the way one has to think about the relation between theory and reality:

When performing on a quantum system a measurement of a quantity A with a physical meaning, one gets an approximation for its value. The thermal interpretation treats the measured value as an approximation not of an eigenvalue of A but of the q-expectation of A, the formal expectation value defined as the trace of the product of A with a density operator describing the state of the system. The approximation error is of the order of the uncertainty σ_A. This postulate is more or less implied by—and hence more cautious than—the traditional postulate that the measured value is an eigenvalue, obtained with the probability given by Born's rule.

This novel postulate of the thermal interpretation remains therefore valid in all cases where the traditional postulates apply. It avoids a number of problems of Born's rule (collected in the Appendix).

For this book, I rearranged, condensed, and augmented the material from a number of preprints (Neumaier [202, 203, 204, 205, 206, 207, 211] and Neumaier & Ghaani Farashahi [212]; see also the exposition at the website Neumaier [200]) such that, after an introductory chapter—explaining the reasons for the book and the main results—the formal "shut-up-and-calculate", probably less controversial part—comes first, the thermal interpretation comes second, and the detailed critique of the tradition (that motivated everything) comes last.

The coherent foundations proposed here in a programmatic way resolve the problems with the traditional presentation of quantum mechanics discussed in the introductory Chapter 1. Part I features the mathematics of quantum physics, a formal core and its development that follows in a purely logical way from basic axioms and definitions that build on it. It gives a coherent, interpretation-independent description of quantum theory. Part II motivates, defines, and develops the thermal interpretation and its implication for the complex of conceptual issues called the measurement problem. Part III is an Appendix containing a detailed critique of Born's rule, a centerpiece of the tradition, of the concepts of states and ensembles, and a comparison to traditional interpretations.

This book is not an introduction to quantum mechanics. Much of the material is intended to be nontechnical, needing only a fairly elementary background. It is aimed at a wide audience that is familiar with some traditional quantum mechanics and basic terms from functional analysis. But another large part of the material is addressed to

experts.[2] There I refer to technical aspects, usually explained in the references given. Thus, where necessary, I draw whatever seems relevant for coherent foundations from functional analysis, quantum mechanics, quantum field theory, and statistical mechanics, while skipping many techniques that are treated in typical textbooks. For the sake of definiteness, the fundamental description of Nature is taken in this book to be given by 4-dimensional relativistic quantum field theory in Minkowski space-time.[3] Since this is ongoing research, I also refer to material that is still unpublished and will appear elsewhere.

For the discussion of questions related to this book, please use the discussion forum Physics Overflow at https://www.physicsoverflow.org. See also my webpage on the thermal interpretation at https://www.mat.univie.ac.at/~neum/physfaq/therm. A list of errata will be maintained there; please report corrections to me at Arnold.Neumaier@univie.ac.at.

I would like to thank Arash Ghaani Farashahi, Waltraud Huyer, Rahel Knöpfel, David Bar Moshe, Mike Mowbray, Karl-Hermann Neeb, Hermann Schichl and Eric Wofsey for useful discussions related to coherent spaces. The material on interpretation benefited from discussions with Hendrik van Hees, Rahel Knöpfel, Mike Mowbray, Paul Pöll, and Francois Ziegler, which are also gratefully appreciated.

The puzzle of making sense of the foundations of quantum physics held my attention for many years. Around 2003, I discovered that group coherent states are for many purposes very useful objects; before, they were—for me—just a facet that physicists (who needed them for quantum optics) studied. In 2007, I realized that apparently all of quantum mechanics and quantum field theory can be profitably cast into this form, and that coherent states may provide better theoretical foundations for quantum mechanics and quantum field theory than the current Fock space approach. Since then I have been putting them bit by bit into the new framework, and always found (after some work) everything nicely fitting. With each new piece in place, I got insights about how to interpret everything, and things got simpler and simpler as I proceeded. Or rather, more and more complicated things became understandable without significantly increasing the complexity of the new picture. Everything became much more transparent and intuitive than the traditional mental picture of quantum physics.

Hints at a possible thermal interpretation of quantum physics go back at least to 1997; see the above quotes by Wetterich and Klauder. The thermal interpretation of quantum physics itself emerged from my foundational 2003 paper Neumaier [194].

2 Nonexperts are advised to simply skip the more advanced passages and continue reading when the discussion becomes again less technical. In particular, Part II does not depend on Chapters 4–6 and Section 7.2, and Part III is independent of Parts I and II.

3 It does not matter whether or not there is a deeper underlying structure, such as that of string theory, in terms of which quantum field theory would be an effective theory only. For simplicity, curved spacetimes are not considered.

It was developed by me in discussions on the newsgroups de.sci.physik, starting in Spring 2004, and in later discussions on PhysicsForums; for the beginnings see Neumaier [196]. A first version of it was fully formalized (without naming the interpretation) in Sections 5.1, 5.4, and Chapter 7 of the 2008 edition of the online book by Neumaier & Westra [214]; see also Sections 8.1, 8.4, and Chapter 10 of the 2011 edition. The term "thermal interpretation" appeared first in a 2010 lecture (Neumaier [197]). Later I created a dedicated website on the topic (Neumaier [198]). A recent view closely related to the thermal interpretation is the 2017 work by Allahverdyan et al. [7].

I trust you will enjoy reading the book!

Vienna, June 14, 2019,
Arnold Neumaier

Foreword

The predictions of quantum mechanics are remarkably accurate, but aspects of the interpretation of quantum mechanics are still not agreed upon.

The use of coherent states as basic tools in quantum mechanics has several advantages. As an example, I cite their role in quantum/classical issues, as illustrated by the quantum action functional given by

$$A_Q = \int \langle \psi(t) | \Big(i\hbar \frac{\partial}{\partial t} - H(P,Q) \Big) | \psi(t) \rangle dt$$

for normalized general Hilbert space vectors $|\psi(t)\rangle$. However, classical observers may be limited to fewer vectors, such as

$$|p(t), q(t)\rangle = e^{-iq(t)P/\hbar} e^{ip(t)Q/\hbar} |any\rangle,$$

which involve moving the system to a new position q or new velocity $v = p/m$. This leads to

$$A_C = \int \langle p(t), q(t) | \Big(i\hbar \frac{\partial}{\partial t} - H(P,Q) \Big) | p(t), q(t) \rangle dt$$

$$= \int \big(p(t)\dot{q}(t) - H(p(t), q(t)) \big) dt + O(\hbar; p(t), q(t)).$$

Observe that this is the classical action functional, with quantum corrections for $\hbar > 0$, as it must be in the real world.

The author's book is full of connections of this sort, and they can certainly help in clarifying quantum mechanics!

August 7, 2019
John R. Klauder

https://doi.org/10.1515/9783110667387-202

Contents

Part II: **The interpretation of quantum physics**

1 Introduction

This chapter sets the informal stage for the subject matter of the book. Section 1.1 presents in concise form what is typically taught as the basics in quantum physics courses around the world. Section 1.2 gives a short overview of the most important interpretations of these basic rules that were spawned by a century-long lack of clarity of the meaning of quantum mechanics. Section 1.3 gives an account of the deplorable tradition of quantum magic that resulted from this lack of clarity. Section 1.4, the final section, gives a preview of the coherent quantum physics proposed in this book as a solution to the problems of interpreting quantum physics.

1.1 The 7 basic rules of quantum mechanics

The 7 basic rules (BR1)–(BR7) given below (taken from NEUMAIER [211]) reflect what is typically taught as the basics in quantum physics courses around the world. They are found in almost all introductory quantum mechanics textbooks.[1] Often they are stated in terms of axioms or postulates, but this is not essential for their practical validity. In some interpretations, some of these rules are not considered fundamental rules, but as empirical or effective rules for practical purposes.

The footnotes contain generalizations of these rules for degenerate eigenvalues, for mixed states, and for measurements not defined by self-adjoint operators, but by POVMs: see Footnote 7 below. These generalizations are necessary to apply quantum mechanics to all situations encountered in practice. The basic rules are carefully formulated so that they are correct as they stand and, at the same time, fully compatible with these generalizations.

(BR1) A **quantum system** is described using a Hilbert space[2] $\mathcal{H}$.

(BR2) A **pure state** of a quantum system is represented by a normalized vector $|\psi\rangle$ in $\mathcal{H}$; state vectors differing only by a phase factor of absolute value 1 represent the same state.[3] In the **position representation**, where the Hilbert space is the space of square integrable functions of a position vector x, $\psi(x)$ is called the **wave function** of the system.

---

1 Among them are: Basdevant 2016; Cohen-Tannoudji, Diu and Laloe 1977; Dirac 1930, 1967; Gasiorowicz 2003; Greiner 2008; Griffiths and Schroeter 2018; Landau and Lifshitz 1958, 1977; Liboff 2003; McIntyre 2012; Messiah 1961; Peebles 1992; Rae and Napolitano 2015; Sakurai 2010; Shankar 2016; Weinberg 2013. Even Ballentine 1998, who rejects rule (BR7), whose process (9.9), as fundamental, derives it in the form of his (9.21) as an effective rule.

2 Often, this Hilbert space is assumed to be separable, that is, to have a countable orthonormal basis.

3 Equivalently, a pure state can be represented by a rank 1 density operator $\rho = |\psi\rangle\langle\psi|$, satisfying $\rho^2 = \rho = \rho^*$ and $\mathrm{Tr}\,\rho = 1$. **Mixed states** are represented by more general (nondempotent) Hermitian density operators of trace 1.

https://doi.org/10.1515/9783110667387-001

(BR3) The time evolution of an isolated quantum system represented by the state vector $|\psi(t)\rangle$ is given by[4]

$$i\hbar\frac{d}{dt}\big|\psi(t)\big\rangle = H\big|\psi(t)\big\rangle,$$

where H is the Hamilton operator and $\hbar$ is Planck's constant. This is the **Schrödinger equation**. This rule is valid in the formulation of quantum mechanics called the Schrödinger picture. There are other, equivalent formulations of the time evolution, especially the Heisenberg picture and the interaction pictures, where time evolution is entirely or partially shifted from the state vector to the operators.

(BR4) An **observable** of a quantum system is represented by a Hermitian operator A with real spectrum[5] acting on a dense subspace of $\mathcal{H}$.

(BR5) The possible **measured values** of a measurement of an observable are the spectral values of the corresponding operator A. In the case of a discrete spectrum, these are the eigenvalues a satisfying $A|a\rangle = a|a\rangle$.

(BR6) Let $\{|a\rangle\}$ be a complete set of (generalized) eigenvectors of the self-adjoint operator A with spectral values a. Let the quantum system be prepared in a state represented by the state vector $|\psi\rangle$. If a measurement of the observable corresponding to A is performed, the probability (density) $p_\psi(a)$ for finding the measured value a is given by

$$p_\psi(a) = \big|\langle a|\psi\rangle\big|^2.$$

This is **Born's rule**, in a formulation that assumes that all eigenvalues are nondegenerate.[6]

(BR7) For successive, nondestructive projective measurements with discrete results,[7] each measurement with measuring value a can be regarded as the **preparation**

---

4 It is equivalent to define the time evolution of an isolated quantum system by

$$\big|\psi(t)\big\rangle = U(t)\big|\psi(0)\big\rangle$$

with the unitary time evolution operator $U(t) = e^{-iHt/\hbar}$. The evolution according to (BR3) is therefore also referred to as **unitary evolution**.

5 Equivalently, A is self-adjoint.

6 In the case of degenerate eigenvalues, let $\{|a,\nu\rangle\}$ be a complete set of (generalized) eigenvectors of A, indexed by ν. The probability $p_\psi(a)$ for finding the measured value a is then given by summing (or integrating) over ν, that is, over the entire a-subspace

$$p_\psi(a) = \sum_\nu \big|\langle a,\nu|\psi\rangle\big|^2.$$

7 The projection postulate is valid only under the assumptions stated, such as passing barriers with holes or slits, polarization filters, and certain other instruments that modify the state of a quantum system passing through it. This (nonunitary, dissipative) change of the state to an eigenstate in the course

of a new state, whose state vector is the corresponding eigenvector $|a\rangle$, to be used for the calculation of subsequent time evolution and further measurements. This is the **von Neumann projection postulate**.

These rules say nothing about the practically very important problem of how to handle a nonisolated quantum system outside of explicit measurement contexts. Hence they are only an approximate guide to the meaning of quantum mechanics in general. Applying the rules in practice requires further assumptions and developments.

1.2 Interpretations of quantum mechanics

My 'orthodoxy' is not identical to that of Bohr, nor to that of Peierls, to mention two especially eminent examples. Hence I must state my definition of 'orthodoxy'.

Kurt Gottfried, 1991 [103, p. 36]

Orthodox QM, I am suggesting, consists of shifting between two different ways of understanding the quantum state according to context: interpreting quantum mechanics realistically in contexts where interference matters, and probabilistically in contexts where it does not. Obviously this is conceptually unsatisfactory (at least on any remotely realist construal of QM) – it is more a description of a practice than it is a stable interpretation. [...] The ad hoc, opportunistic approach that physics takes to the interpretation of the quantum state, and the lack, in physical practice, of a clear and unequivocal understanding of the state – this is the quantum measurement problem.

David Wallace, 2016 [292, p. 22, p. 24]

Not further discussing the foundations of quantum mechanics beyond this is called **shut-up-and-calculate**. It is the mode of working sufficient for all who do not want to delve into often highly disputed foundational (and partly philosophical) problems. However, the above-mentioned rules are often considered conceptually unsatisfactory because they introduce not well-defined terms "probability", "measurement", and "observer" to define these basic rules, whereas in principle one expects that at least measurement and observation can be regarded as quantum mechanical processes or interactions, which follow the same fundamental rules and do not play any special role. The associated issues are treated in different ways by different **interpretations of quantum mechanics**.

In the **Copenhagen interpretation** (also called **standard interpretation** or **orthodox interpretation**; terminology and interpretation details vary), the above rules

---

of a projective measurement is often referred to as "state reduction" or "collapse of the wave function" or "reduction of the wave packet". Note that there is no direct conflict with the unitary evolution in (BR3), since during a measurement a system is never isolated.

In other cases, the prepared state may be quite different; see the discussion in Landau & Lifschitz [171, Section 7]. The most general kind of quantum measurement and the resulting prepared state is described by so-called positive operator valued measures (POVMs).

are simply operational rules that work in practice. The state vector is a tool that one uses to calculate the probabilities of measurement outcomes, and one is agnostic about whether the state vector represents any object that exists in reality. Rules (BR6) and (BR7) apply only when a measurement has occurred. Thus, unlike in classical physics, it is not enough to specify the initial conditions of the state, and let the state evolve. One must also specify when a measurement has occurred: Generally, a measurement is understood to have occurred when a definite (irreversible, that is, nonunitary) measurement result or outcome has been obtained. For example, the observer records a mark on a screen. (However, passing a Stern–Gerlach magnet—which in modern terminology is a premeasurement only—is frequently, but inaccurately, considered to be a measurement, although it is described by a unitary process, where even in principle no measurement result becomes available.)

A noteworthy aspect of the standard interpretation is that the state vector cannot represent the whole universe, but must exclude an observer or measuring apparatus that decides when a measurement has occurred; this is the so-called **Heisenberg cut** between the quantum and the classical world. To date, this has not been a problem in making successful experimental predictions, so practitioners are often satisfied with the quantum formalism and the standard interpretation.

However, many have suggested that there is a conceptual problem with the standard interpretation because the whole universe presumably obeys laws of physics. So there should be laws of physics that describe the whole universe, without any need to exclude any observer or measurement apparatus from the quantitative description. Then one must be able to derive the rules (BR5)–(BR7) for measuring subsystems of the universe from the dynamics of the universe. The problem of how to do this is called the **measurement problem**. A related problem, the problem of the emergence of a classical macroscopic world from the microscopic quantum description, is often considered as essentially solved by **decoherence**.

To solve the measurement problem, other interpretations of the quantum formalism or theories have been proposed. These alternative interpretations or theories are based on different postulates than those of the standard interpretation, but seek to explain why the standard interpretation has been so successful (for example, by deriving the rules of the standard interpretation from other postulates). The major alternative interpretations or theories that have been proposed include **Everett's relative state interpretation** (or **many worlds interpretation**), the **ensemble interpretation** (or **minimal statistical interpretation**), the **transactional interpretation**, and the **consistent histories interpretation**.

Still other interpretations (for example, **Bohmian mechanics**, **Ghirardi–Rimini–Weber theory**, the **cellular automaton interpretation**, and the **thermal interpretation**) modify one or more of the 7 basic rules, and only strive to derive the latter in some approximation for all practical purposes (FAPP). In particular, rule (BR7) cannot be fundamental if one wants to interpret the state vector $|\psi\rangle$ in an ontic way, that is,

as some direct and 'faithful' representation of 'externally existing reality' independent from any observer, observation or measurement.

An interpretation of quantum mechanics relates the formalism to the actual informal practice of using quantum mechanics in our scientific culture. It must show that there is a consistent relation between theory and practice, but to show this, it may use objects familiar from our culture without having to explain their working.

The minimal statistical interpretation does this for predicting the outcome of experiments. It is silent about the interpretation of quantum mechanics in the absence of measurements, and in particular about the interpretation of quantum physics applied to the far past before experiments were possible. This is a serious gap, but it is consistent, just incomplete (as it should be for a "minimal" interpretation).

The Copenhagen interpretation that claims (in its most radical version) that nothing can be asserted in the absence of a measurement is also consistent. But this sounds like the concept that a tree fallen in the wood has fallen only after someone has seen it, and is part of the reason (see Section 1.3 below) why quantum mechanics comes across as somewhat strange. In a many-world interpretation anything goes, and at not even specifiable times, the world splits and splits completely unnoticed by us. This is already science fiction by conception.

The other known interpretations are either variations of the above or require additional, in principle, unobservable and hence fictional stuff. Thus, none of the traditional interpretations is satisfactory.

According to the thermal interpretation of quantum physics featured in Part II of this book, Nature existed before human minds existed and observed it. Nature has now and had then objective properties comprehended and described locally by q-expectations of quantum fields and nonlocally by more complicated q-expectations. Perceptions (including experiments and measurements) and the resulting knowledge only provide approximations to these objective properties. People have better approximations about precisely those aspects about which they are more knowledgeable. The laws and symmetries of standard quantum field theory are taken to apply exactly to Nature, though we only approximately know the details of the field content, the detailed interactions, and the detailed state of the universe. We can only explain part of the history and predict part of the future since our knowledge and understanding of the true state of the universe is limited.

1.3 Quantum magic

Die Wahrscheinlichkeitsinterpretation (insbesondere für das spontane Auftreten von Partikeleigenschaften) wird allen Physikstudenten als unumstößliches Dogma ins Gehirn gebrannt. Sie ist für viele Zwecke natürlich gerechtfertigt, beschreibt jedoch nur die halbe Wahrheit über die Wellenfunktion und überläßt die Anwendung der dabei zu benutzenden statistischen Regeln weitgehend der situationsbedingten Intuition.

Dieter Zeh, 2012 [314, p. 47]

> *I consider it to be an intellectual scandal that, nearly one hundred years after the discovery of matrix mechanics by Heisenberg, Born, Jordan and Dirac, many or most professional physicists – experimentalists and theorists alike – admit to be confused about the deeper meaning of Quantum Mechanics (QM), or are trying to evade taking a clear standpoint by resorting to agnosticism or to overly abstract formulations of QM that often only add to the confusion.*
>
> Jürg Fröhlich, 2019 [92, p. 1]

Traditionally, those learning quantum theory are expected to abandon classical thinking and to learn thinking in a quantum mechanical framework completely different from that of classical mechanics. Though students widely differ in the order in which this happens, sooner or later, most of them are introduced to the items mentioned in the following caricature:

- Typically, they are introduced to quantum mechanics by Planck's explanation of black body radiation and the Bohr–Sommerfeld quantization rules explaining the spectral lines for the hydrogen atom, firmly establishing that Nature is quantized.
- Then they are told that Bohr's view is obsolete, and that it was just a happy (or even misleading) coincidence that the old quantum theory worked for hydrogen.
- Therefore, they are next acquainted with wave functions on configuration space, their inner product, and the resulting Hilbert spaces of square integrable wave functions.
- But almost immediately, unnormalizable bras and kets are used that do not belong to the Hilbert space.
- They are told with little intuitive guidance (except for a vague postulated correspondence principle that cannot be made to work in many cases) that in quantum mechanics, observables are replaced by Hermitian operators on this Hilbert space.
- Later they may (or may not) learn that many of these operators are not even defined on the Hilbert space, but only on a subspace.
- Then they must learn that between measurements, position and momentum – and hence well-defined paths – do not exist, but that when measured, they miraculously get random values.
- They are taught the connection to classical physics by establishing the Ehrenfest theorem for expectation values that obey, approximately, classical laws. Miraculously, the system has at all times a well-defined mean path, even when not measured.
- They must swallow a mysterious law defining the distribution of these random values, called Born's rule. It is justified by the remark that it is proved by the Stern–Gerlach experiment. But Born's rule is claimed to hold for all conceivable quantum measurements, although this experiment neither demonstrates the measurement of position, nor of momentum or other important quantities.
- No explanation is given how the Stern–Gerlach screen can possibly find out that a particle—without having position or momentum—arrives to be measured.

The Stern–Gerlach device and all of quantum mechanics begins to look like magic.[8]

- They are made familiar with the postulated collapse of the wave function that prepares the system in an eigenstate of the measured observable.
- But for measuring position or momentum, these eigenstates do not exist since they are unnormalizable.
- As a result, classical and quantum physics appear like totally separated realms with totally different concepts and tools, connected only by a rough-and-ready notion of correspondence that is ambiguous and never made precise, but works in a few key cases (and always with liberally enough usage).
- After considerable time, when they have some experience with spectral calculations, they learn how to use group theory (or, for those with only little algebra background, spherical harmonics—rotation group representation tools in disguise) to determine the spectrum for hydrogen. Miraculously, the results are identical with those obtained by Bohr, whose approach was earlier declared to be obsolete.
- At a far later stage, they meet (if at all) coherent states for describing laser light, or as a tool for a semiclassical understanding of the harmonic oscillator. Miraculously, a coherent state happens to perform under the quantum dynamics exact classical oscillations.
- Only few students will also meet coherent states for the hydrogen atom, the Berry phase, Maslov indices, and the accompanying theory of geometric quantization, which gives the (slightly corrected) Bohr–Sommerfeld rules for the spectrum a very respectable place in the quantum theory of exactly solvable systems, even today relevant for semiclassical approximations.
- Even fewer students realize that this implies that, after all, classical mechanics and quantum mechanics are not that far apart. A development of quantum mechanics emphasizing the closeness of classical mechanics and quantum mechanics can be found in the online book by Neumaier & Westra [214].

Why does the conventional curriculum lead to such a strange state of affairs? Perhaps this is the case because tradition builds the quantum edifice on a time-honored foundation, which accounts for essentially all experimental facts, but takes a "shut-up-and-calculate" attitude with respect to the interpretation of the foundations. The traditional presentation of quantum physics is clearly adequate for prediction, but seems not to be suitable for an adequate understanding.

Another reason might be that the weirdness in quantum mechanics seems to play an important, entertaining social role in the communication of physics. In the quan-

8 In quantum information theory, there is another, formal meaning of the terms "magic", which has nothing to do with this informal magic.

tum domain, there are very popular **quantum magicians** who are at the same time very experienced quantum physics practitioners specializing in quantum optics. They entertain the world with well-prepared quantum weirdness. They like to create for their audience the impression that important parts of quantum mechanics are weird. And the general public loves it!

This is common to magicians in any field, and not specific to quantum mechanics. People very experienced in a particular area of real life can easily trick those who do not understand the corresponding matter well enough, leading them into believing that seemingly impossible things can happen. This is true in the classical domain, amply documented by magic tricks, where really weird things happen, such as rabbits being pulled out of empty hats.

The art of a magician consists in studying particular potentially weird aspects of Nature and presenting them in a context that emphasizes the weirdness. Part of the art consists of remaining silent about the true reasons why things work rationally, since then the weirdness is gone, and with it the entertainment value.

Judging by its social impact, quantum weirdness will never go away as long as highly reputed scientists are willing to play the role of a quantum magician. But only the presentation makes quantum mechanics appear weird. It is fully rational to the mind sufficiently trained in mathematics and theoretical physics.

Does quantum mechanics have to be weird? It sells much better to the general public if it is presented that way, and there is a long history of proceeding that way. But, in fact, it is an obstacle for everyone who wants to truly understand quantum mechanics, and to physics students who have to unlearn what they were told as laypersons.

When presented in the right way, quantum mechanics is not at all weird, but very close to classical mechanics. Much of the weirdness comes from forcing quantum mechanics into the straightjacket of a particle picture. The particle picture breaks down completely in the subatomic domain, as witnessed by the many weird things such a view leads to. On the other hand, the field picture remains valid at all length and time scales.

1.4 Coherent quantum physics

Are coherent states the natural language of quantum theory?

John Klauder, 1986 [159, title]

In 1927, when the Copenhagen interpretation (the informal agreement on the interpretation reached at the 1927 Como and Solvay conferences) was forged, its main purpose was to reconcile the then new quantum formalism with the experimental evidence available at that time. Apart from the Stern–Gerlach experiment, the evidence consisted exclusively of (i) the observation of spectra of atoms and molecules, and (ii) the

need to reconcile the quantum description of the invisible microscopic details with the classical description of the macroscopic world.

Section 2.3 (below) shows that the same evidence is naturally explained by the thermal interpretation. Indeed, with a little more work and imagination, Paul Ehrenfest, whose paper [76] appeared in 1927, could have easily found and justified this interpretation.

Coherent quantum physics removes the radical split between classical mechanics and quantum mechanics. That this might be feasible is already suggested by the history of coherent states. (For the early history of coherent states see, for example, Nieto [219].) In 1926, at the very beginning of modern quantum physics, coherent states were used by Schrödinger [269] to demonstrate the closeness of classical and quantum mechanical descriptions of a physical system. Schrödinger discussed the main properties of the coherent states today known as Glauber coherent states. He did not call them coherent states, a notion coined in 1963 by Glauber [98]. Today the term **coherent state** denotes a variety of (in detail very different) collections of states displaying simultaneously a classical and a quantum character.

In this book, the fundamental importance of coherent states is emphasized by defining a **coherent quantum physics**, based on the concept of coherence in various forms, thereby rebuilding from scratch the foundations of quantum physics. Summarizing the vision in the shortest terms, we may say:

- Coherent quantum physics is physics in terms of a coherent space (see Chapter 5) consisting of a line bundle over a classical phase space and an appropriate "coherent product" characterizing the physical properties of a quantum system.
- The kinematical structure of quantum physics and the meaning of the fundamental quantum observables are given (in Chapter 6) by the symmetries of this coherent space, their infinitesimal generators, and associated operators on the quantum space of the coherent space.
- The connection of quantum physics to experiment is given through the thermal interpretation, defined in Section 9.2. The dynamics of quantum physics is given (for isolated systems) by the Ehrenfest equations for q-expectations (Section 2.2).

Coherent spaces reconcile the old (semiclassical, Bohr-style) thinking with the requirements of the new (operator-based) quantum physics. They become the foundation on which a better, coherent quantum physics is built. Mathematically, these foundations are equivalent to the traditional Hilbert space approach. But conceptually, these foundations begin with what is common between the classical and the quantum world.

It is a linguistic coincidence that the same word "coherence" fits several different contexts that come together in coherent quantum physics:

- spatio-temporal coherence, the meaning of the term in "coherent states";
- logical coherence, referring to mathematically sound foundations;

- intuitive coherence, implying that concepts make holistic sense to the intellect; and
- coherence as harmony, the meaning of the term in "coherent configurations" and "coherent algebras", concepts from the combinatorics of symmetry.[9]

In the literature, one usually finds coherent states discussed just for themselves, or in the context of the classical limit. However, they are also a powerful instrument in other respects. The reason is that they have both a good intuitive semiclassical interpretation and give good access to the whole Hilbert space (and beyond). Indeed, coherent quantum physics turns coherent states into the fundamental tool for studying quantum physics.

In Section 6.6, we rephrase the formal properties of q-expectations and uncertainty in a more abstract, slightly generalized setting, to emphasize the essential mathematical features and the close analogy between classical and quantum physics. We use the coherent action principle (the Dirac–Frenkel variational procedure applied to coherent states) to show that in coarse-grained approximations that only track a number of relevant variables, quantum mechanics exhibits chaotic behavior. According to the thermal interpretation, featured in Part II, this is responsible for the probabilistic aspects of quantum mechanics.

1.5 Overview to Part I

Part I is concerned with an exposition of the purely formal, mathematical part of quantum physics, as far as necessary for a conceptual understanding of the foundations.

Chapter 2 discusses the formal core of quantum physics. It presents, in particular, the Ehrenfest picture of quantum mechanics, which expresses everything in terms of q-expectations, thus giving the latter a prominent place in the theory. This, particu-

9 Coherent configurations and the associated coherent algebras were introduced in pure mathematics around 1970 (by Higman [130, 131]), completely independent of physical considerations. Special classes of coherent configurations called **association schemes** and **distance-regular graphs** are very well-studied, and many interesting examples are known in detail.
Coherent configurations are, in a sense made precise in Neumaier [208], closely related to a finite variant of coherent states. Just as semisimple Lie groups act as symmetry groups of associated Riemannian symmetric spaces, and their representation theory leads to Perelomov coherent states, so most finite simple groups act as symmetry groups of associated distance regular graphs. The latter is recorded in the book by Brouwer et al. [55], of which I am a coauthor. The book appeared just about the time when, for completely different reasons, I turned to seriously study quantum physics.
Only much later, I realized the extent of the connection of that work to coherent states. The two subjects (quantum physics and the combinatorics of symmetry) met in the past only in one area, the study of **symmetric, informationally complete, positive operator valued measures** (**SIC-POVM**s); see, for example, [308]. Today's main open problem in the study of SIC-POVMs is **Zauner's conjecture**, which dates back to the 1999 Ph. D. thesis of Zauner [313], written under my supervision.

larly, allows us to derive classical physics from quantum physics in the approximation in which uncertainties can be neglected, and to derive the basic spectroscopic consequences of quantum mechanics in terms of resonance phenomena.

Chapter 3 precisely defines the basic notion of uncertainty, and shows how probability and statistics arise from this notion of uncertainty together with the weak law of large numbers.

Chapter 4 discusses Euclidean spaces—abstractions of the spaces of Schwartz functions that form a common dense domain of the unbounded operators algebras (for example, that generated by position and momentum operators) that play a critical role in practical quantum physics. It also prepares the stage for Chapter 5, which gives rigorous definitions of the basic concepts and results on coherent spaces, without attempting to be comprehensive. It focuses on introducing the concept of symmetries of coherent spaces relevant for quantization.

Chapter 6 gives a general outline of a coherent quantum physics, telling the main points of the story with as few formulas and conceptual details as justifiable. A quantization procedure based on Neumaier & Ghaani Farashahi [212] leads to quantum dynamics, which in special (completely integrable) situations can be solved in closed form in terms of classical motions on the underlying coherent space, if the latter has a compatible manifold structure. Spectral issues can, in favorable cases, be handled in terms of dynamical Lie algebras. Close relations to concepts from geometric quantization and Kähler manifolds are pointed out.

Chapter 7 defines the meaning of the notion of a field in the abstract setting of Section 6.6, and shows how coherent spaces may be used to define relativistic quantum field theories. This chapter also contains the essentials of quantum statistical mechanics and its important concept of coarse-graining. This provides the link between microscopic physics and the world of ordinary experience.

1.6 Overview to Part II

Part II is devoted to a coherent interpretation of quantum physics called the **thermal interpretation**. We begin in Chapter 8 with a detailed analysis of the requirements for good foundations of a physical theory. Just as in geometry and number theory, the concepts of a mature theory should be essentially self-explaining, in a sense explained there. Chapter 9 then gives a concise exposition of the thermal interpretation, its advantages, and some open problems. The analogy to classical physics is striking!

According to the standard interpretation of classical physics, particles exist in classical physics. States define their objective properties. These are given by the exact positions and momenta of the particles, some of which can be approximately measured. The basic dynamics, given by Newton's equations of motion, has the structure of deterministic Hamiltonian mechanics. From a fundamental point of view, fields are (as in classical continuum mechanics) only coarse-grained approximate concepts.

According to the thermal interpretation of quantum physics, fields exist in quantum physics. States define their objective properties. These are given by the exact q-expectations of the fields and their appropriately normally ordered or time-ordered products, some of which can be approximately measured. The basic dynamics, given by the Ehrenfest equations, has the structure of deterministic Hamiltonian mechanics. From a fundamental point of view, particles are (as in classical geometric optics) only coarse-grained approximate concepts.

Experimental physics is, in both cases, about how to do the measurements, and under which conditions which measurements are how accurate. This is achieved using the standard theory based upon three ingredients: the formal core of quantum mechanics, the respective foundations, and Callen's criterion that, operationally, a system is in a given state if its properties are consistently described by the theory for this state.

Chapters 10 and 11 discuss the measurement problem from the point of view of the thermal interpretation. It is shown how Born's rule, traditionally taken to be fundamental, arises in measurement situations from uncontrollable influences of the environment. The discrete nature of certain quantum phenomena is explained by the switch-like dissipative bistability of the coarse-grained approximations used to model open systems. Chapter 12 discusses the emergence of particles as asymptotic entities from the more fundamental quantum field point of view. Chapter 13 discusses several model experiments in the light of the thermal interpretation.

1.7 Overview to Part III

Part III is an appendix featuring a critique of the tradition of quantum physics in its mainstream interpretation (that is, treating pure states and probability as primitives, without reference to hidden variables, and without modifications of the quantum laws), at the same time relating it to the thermal interpretation.

This is achieved by cleanly separating in Chapter 14 a concise version of the (universally accepted) formal core of quantum physics (described in Section 2.1) from the (controversial) interpretation issues. The latter are primarily related to measurement, but also to questions of existence and of the meaning of basic concepts like "state" and "particle".

The bridge between the formal core and measurement is traditionally taken to be Born's rule, assumed to be exact. It is argued in Chapter 14 why this assumption cannot be maintained. Thus, Born's rule must be considered as a scientific law with a restricted domain of validity.

In Chapter 15, the concepts of states and ensembles are critically reviewed. Among others, it is shown that if the state of every composite quantum system contains all information that can be known about a system, states cannot be pure states.

A multitude of other interpretations of quantum mechanics exist; most of them in several variants. The final Chapter 16 discusses some of the traditional interpretations in the light of the thermal interpretation. As we shall see, the mainstream interpretations may be regarded as partial versions of the thermal interpretation. In particular, certain puzzling features of both the Copenhagen interpretation and the statistical interpretation get their explanation through the thermal interpretation of quantum physics. It is shown that both the Copenhagen interpretation and the statistical interpretation have a restricted domain of validity—the statistics of few particle scattering events, where they may be viewed as limiting cases of the thermal interpretation.

2 Basic quantum physics

Quantum physics consists of a **formal core** that is universally agreed upon (basically being a piece of mathematics with suggestive names for concepts matching related concepts of experimental culture) and an interpretational halo (trying to make this relation more precise) that remains highly disputed even after more than 90 years of modern quantum physics. The latter is the subject of the **interpretation of quantum mechanics** (see Part II), where many interpretations coexist and compete for the attention of those interested in what quantum physics really means.

In this chapter, an axiomatic introduction to the undisputed formal core of quantum physics is given. The axioms described present nonrelativistic quantum statistical mechanics in the Schrödinger picture. The relativistic case is outside the scope of these axioms, as it must be treated by quantum field theory; see Chapter 7.

As in any axiomatic setting (necessary for a formal discipline), there are a number of different but equivalent sets of axioms or postulates that can be used to define formal quantum physics. Since they are equivalent, their choice is a matter of convenience. The choice presented here is a formulation featuring three aspects: It emphasises the similarity of quantum mechanics and classical mechanics, gives most direct access to statistical mechanics, and is free from allusions to measurement.

The reason for the first is that this similarity serves as a guide to meaning. From the axiomatic setting for quantum mechanics presented below, classical mechanics (in the Koopman setting discussed in Section 7.10) is obtained by simply restricting the operators to be diagonal, so that all operations happen pointwise on the diagonal elements. Thus, multiplication is commutative, and one can identify operators and functions. In particular, the density operator degenerates into a probability density. Therefore, the quantum case appears as a generalization of the classical case allowing for noncommutativity, so that both q-observables and the density operator are (usually infinite-dimensional) operators.

The reason for the second is that statistical mechanics is the main tool for applications of quantum physics to the macroscopic systems we are familiar with.

The reason for the third is that realistic measurements constitute a complex process involving macroscopic detectors, hence should be explained by quantum statistical mechanics rather than be part of the axiomatic foundations themselves. (This is in marked contrast to other foundations, and distinguishes the present system of axioms.)

2.1 Axioms for the formal core of quantum physics

In full generality (that is, without the simplifications presented in the 7 basic rules of Section 1.1), quantum physics is governed by the six axioms (A1)–(A6)[1] that follow.

https://doi.org/10.1515/9783110667387-002

Note that all axioms are basis-independent. $\hbar$ is Planck's constant, and is often set to 1.

(A1) A generic system (for example, a "hydrogen molecule") is defined by specifying a (complex) Hilbert space $\overline{\mathbb{H}}$ and a self-adjoint[2] linear operator H, called the **Hamiltonian** (or the **internal energy**), mapping a dense subspace $\mathbb{H}$ of $\overline{\mathbb{H}}$ (usually a nuclear space) into itself.

(A2) A particular system (for example, "the ion in the ion trap on this particular desk") is characterized by its **state** $\rho(t)$ at every[3] time t in a time interval. Here $\rho(t)$ is a Hermitian, positive semidefinite, linear trace class operator on $\mathbb{H}$, satisfying at all times the normalization condition

$$\operatorname{Tr}\rho(t) = 1.$$

Here Tr denotes the trace.

(A3) A system is called **closed** (or **isolated**) in a time interval $[t_1, t_2]$ if it satisfies the evolution equation

$$\frac{d}{dt}\rho(t) = \frac{i}{\hbar}[\rho(t), H] \quad \text{for } t \in [t_1, t_2], \tag{2.1}$$

and **open** otherwise. If nothing else is apparent from the context, a system is assumed to be closed.

(A4) Besides the internal energy H, certain other densely defined, self-adjoint operators on $\mathbb{H}$, or vectors of such operators, are distinguished as **quantum observables**, short **q-observables**. (For example, the q-observables for a system of N distinguishable particles conventionally include for each particle several 3-dimensional vectors: the **position** $\mathbf{q}^a$, **momentum** $\mathbf{p}^a$, **orbital angular momentum** $\mathbf{L}^a$, and the **spin vector** (or **Bloch vector**) $\mathbf{S}^a$ of the particle with label a. If $\mathbf{u}$ is a 3-vector of unit length, then $\mathbf{u}\cdot\mathbf{p}^a$, $\mathbf{u}\cdot\mathbf{L}^a$, and $\mathbf{u}\cdot\mathbf{S}^a$ define the momentum, orbital angular momentum, and spin of particle a in direction $\mathbf{u}$.)

(A5) For any particular system, and for every vector X of self-adjoint q-observables with commuting components, one associates a time-dependent monotone linear functional $\langle\cdot\rangle_t$, defining the **q-expectation**

$$\langle f(X)\rangle_t := \operatorname{Tr}\rho(t)f(X)$$

---

1 The statements of the axioms contain in parentheses some additional explanations that, strictly speaking, are not part of the axioms, but make them more easily intelligible. The list of examples given only has illustrative character and is far from being exhaustive.

2 A linear operator H is **self-adjoint** iff it is Hermitian, $H^* = H$, and its spectrum is real. Hermitian trace class operators are always self-adjoint. The Hille–Yosida theorem says that e^{iX} exists and is unitary if and only if X is self-adjoint; see Thirring [286] or Reed & Simon [246].

3 Typical systems considered in practice are defined only in a bounded time interval $[t_1, t_2]$. However, it is common (and for scattering processes necessary) to idealize and allow arbitrary times $t \in \mathbb{R}$, where $\mathbb{R}$ denotes the set of real numbers.

of bounded continuous functions $f(X)$ at time t. By WHITTLE [302], this is equivalent to a multivariate probability measure $d\mu_t(X)$ on a suitable σ-algebra over the spectrum $\operatorname{Spec} X$ of X, defined by

$$\int d\mu_t(X) f(X) := \operatorname{Tr} \rho(t) f(X) = \langle f(X) \rangle_t.$$

(This σ-algebra is uniquely determined and defines **q-probabilities**.)

(A6) Quantum mechanical predictions consist of predicting **properties** (typically q-expectations or conditional q-probabilities) of the measures defined in Axiom (A5), given reasonable assumptions about the state, such as ground state and equilibrium state.

Axiom (A6) specifies that the formal content of quantum physics is covered exactly by what can be deduced from Axioms (A1)–(A5), without anything else added – except for restrictions defining the specific nature of the states and q-observables, for example, specifying commutation or anticommutation relations between some of the distinguished q-observables. Thus, Axiom (A6) says that Axioms (A1)–(A5) are complete.

The description of a particular closed system is therefore given by the specification of a particular Hilbert space (in Axiom (A1)), the specification of the q-observables (in Axiom (A4)), and the specification of conditions singling out a particular class of states (in Axiom (A6)). (The description of an open system involves, in addition, the specification of details of the dynamical law.)

Given this, everything predictable in principle about the system is determined by the theory, and hence is predicted by the theory.

At the level of the formal core, q-expectations and all other concepts introduced are only calculational tools that enables one to predict numerical values for theoretical objects with suggestive names. Their precise relation to experimental reality is not specified by the formal core. These relations are the subject matter of the interpretation of quantum mechanics, whose discussion is deferred to Part II.

2.2 The Ehrenfest picture of quantum mechanics

As first observed in 1925 by DIRAC [68], classical mechanics and quantum mechanics look very similar when written in terms of the Poisson bracket.

Quantities are represented in classical mechanics by functions from a space of suitable smooth phase space functions $A(p,q)$, and in quantum mechanics by linear operators A on a suitable Euclidean space $\mathbb{H}$. We define the **classical Lie product**

$$A \angle B := \{B, A\} = \partial_p A \partial_q B - \partial_p B \partial_q A \tag{2.2}$$

(read $\angle$ as "Lie") of classical quantities A, B, and the **quantum Lie product**

$$A \angle B := \frac{i}{\hbar}[A, B] = \frac{i}{\hbar}(AB - BA) \tag{2.3}$$

of quantum mechanical quantities A, B. This infix notation is much more comfortable than the customary bracket notation. In both cases, it is easy to verify anticommutativity,

$$A \angle B = -B \angle A,$$

the product rule

$$A \angle BC = (A \angle B)C + B(A \angle C),$$

and the Jacobi relation

$$A \angle (B \angle C) = (A \angle B) \angle C + B \angle (A \angle C).$$

This shows that $\angle$ turns the space of quantities into a Lie algebra. It also shows that the application of $a \angle$ to a quantity behaves like differentiation.

We write $\int$ both for the Liouville integral

$$\int A := \int A(p, q)\, dp\, dq \tag{2.4}$$

of a classical quantity A and for the trace

$$\int A := \operatorname{Tr} A \tag{2.5}$$

of a quantum mechanical quantity A. With this notation, it is easy to verify the invariance under infinitesimal canonical transformations,

$$\int A \angle B = 0,$$

from which one finds the integration by parts formula

$$\int (A \angle B)C = \int A(B \angle C).$$

In the general theory, for which we refer to Neumaier [194, 195], these rules are part of a system of axioms for **Euclidean Poisson algebras**, which allows one to develop everything without reference to either the classical or the quantum case.

Quantities and linear functionals are, in general, time-dependent; so we write $\langle f \rangle_t$ for the q-expectation of $f(t)$ at time t. In maximal generality, a **q-expectation** is written in the form[4]

$$\langle A \rangle_t := \int \rho(t) A(t), \tag{2.6}$$

4 In the quantum case, this is the familiar formula $\langle A \rangle_t = \operatorname{Tr} \rho(t) A(t)$.

where $A(t)$ is an arbitrary time-dependent quantity and $\rho(t)$ a time-dependent density quantity, a nonnegative Hermitian operator normalized by

$$\int \rho = 1.$$

Their dynamics is given by[5]

$$\dot{A}(t) = H_1(t) \angle A(t) \tag{2.7}$$

for quantities A, and by

$$\dot{\rho}(t) = \rho(t) \angle H_2(t) \tag{2.8}$$

for the density operator ρ. Note the different treatment of quantities and the density operator! Here $H_1(t)$ and $H_2(t)$ are arbitrary time-dependent expressions without independent physical meaning; they need not satisfy the differential equations (2.7) or (2.8). Integrating (2.8) shows that $\int \rho$ is time independent, so that the dynamics is consistent with the normalization of ρ.

As a consequence of the dynamical assumptions (2.7)–(2.8), the q-expectations (2.6) have a deterministic dynamics, given in terms of $H = H_1 + H_2$ by

$$\frac{d}{dt}\langle A\rangle_t = \langle H \angle A\rangle_t. \tag{2.9}$$

We call (2.9) the **Ehrenfest equation** since the special case of this equation, where A is a position or momentum variable and $H = \frac{p^2}{2m} + V(q)$, is the sum of kinetic and potential energy was found in 1927 by Ehrenfest [76]. Due to the canonical commutation rules, we have

$$\frac{d}{dt}\langle q\rangle_t = \langle H \angle q\rangle_t = \frac{\langle p\rangle_t}{m}, \quad \frac{d}{dt}\langle p\rangle_t = \langle H \angle p\rangle_t = \big\langle -\nabla V(q)\big\rangle_t. \tag{2.10}$$

---

5 More generally, if z is a vector of quantities satisfying (2.7), quantities given by expressions $A(t) = A(z(t), t)$ with an explicit time dependence satisfy (instead of (2.7)) a differential equation of the form

$$\dot{A}(t) = i\big[H_1(t), A(t)\big] + \partial_t A\big(z(t), t\big).$$

This follows easily from (2.7) and the chain rule. The generality gained is only apparent since the numbers $\langle A(z(t), t)\rangle_t$ are expressible in terms of canonical ones: In terms of a Fourier expansion

$$A\big(z(t), t\big) = \int d\omega e^{i\omega t} A_\omega\big(z(t)\big),$$

we see that

$$\big\langle A\big(z(t), t\big)\big\rangle_t = \int d\omega e^{i\omega t} \big\langle A_\omega\big(z(t)\big)\big\rangle_t,$$

and the $A_\omega(z(t))$ are canonical quantities, respecting (2.7). This allows us to limit the main text to the case where A has no explicit t-dependence.

Note that the Ehrenfest equation does not involve notions of reality or measurement, hence belongs to the formal core of quantum mechanics, and is valid independent of issues of interpretation.

The product rule implies that $\frac{d}{dt}\langle A\rangle_t$ only depends on the sum $H = H_1 + H_2$, not on H_1 and H_2 separately. Thus, there is a kind of gauge freedom in specifying the dynamics, which can be fixed by choosing either H_1 or H_2 arbitrarily. Fixing $H_1 = 0$ (so that $H_2 = H$) makes all quantities A time-independent and defines the **Schrödinger picture**. Fixing H_1 as a reference Hamiltonian without interactions (so that $H_2 = V := H - H_1$ is the **interaction**) defines the **interaction picture**. Fixing $H_2 = 0$ (so that $H_1 = H$) makes the density operator ρ time-independent, and defines the **Heisenberg picture**. In the Heisenberg picture, one finds that

$$\big\langle A(u)\big\rangle_s = \big\langle A(u+s-t)\big\rangle_t \tag{2.11}$$

for arbitrary times s, t, u.

The Schrödinger picture is fully compatible with the formal core of quantum physics, comprising the Axioms (A1)–(A6) discussed in Section 2.1. In particular, in the Schrödinger picture, the von Neumann equation

$$\frac{d}{dt}\rho(t) = \frac{i}{\hbar}[\rho(t), H] \quad \text{for } t \in [t_1, t_2] \tag{2.12}$$

holds for closed systems, giving a deterministic dynamics for the density operator.

In place of the traditional Heisenberg, Schrödinger, and interaction pictures, one can also consider another equivalent picture, in which only q-expectations figure as dynamical variables. The name **Ehrenfest picture** is suggestive since, for the standard multiparticle Hamiltonian and $f = p, q$, this reduces to the Ehrenfest equation. In terms of the Lie bracket on q-expectations, defined by the formula

$$\langle A\rangle \angle \langle B\rangle := \langle A \angle B\rangle, \tag{2.13}$$

the family of q-expectations[6] becomes a Lie algebra $\mathbb{L}$, and the Ehrenfest equation (2.9) becomes

$$\frac{d}{dt}\langle A\rangle = \langle H\rangle \angle \langle A\rangle. \tag{2.14}$$

Equation (2.14) is quite remarkable, as it is manifestly independent of how H is split and how the time-dependent expectation is expressed as (2.6).

The Ehrenfest picture gives a complete picture of the (classical or) quantum kinematics and a deterministic dynamics for the q-expectations that is equivalent to the Schrödinger picture, the Heisenberg picture, and the interaction pictures.

6 Strictly speaking, the family of labels for q-expectations (since q-expectations themselves are just complex numbers).

It is interesting to interpret the above in terms of Hamiltonian dynamics on Poisson manifolds. An extensive discussion of classical Hamiltonian dynamics on Poisson manifolds, in particular using Lie–Poisson brackets, and its application to rigid rotors and fluid dynamical systems, is given in Marsden & Ratiu [181]. A **Poisson manifold** is a smooth manifold together with a Lie product on $\mathbb{E} = C^\infty(M)$ that turns $\mathbb{E}$ into a commutative Poisson algebra. Associated with the Lie algebra $\mathbb{L}$ of q-expectations is the manifold $\mathbb{L}^*$ of continuous linear functionals on $\mathbb{L}$. On $\mathbb{E} = C^\infty(\mathbb{L}^*)$, a Lie product is given by the classical Lie–Poisson bracket, which canonically extends the formula (2.13) to smooth functions of q-expectations. This turns $\mathbb{E}$ into a commutative Poisson algebra, hence $\mathbb{L}^*$ into a Poisson manifold.[7] In these terms, the Ehrenfest picture of *quantum* mechanics is just *classical* (but nonsymplectic) Hamiltonian dynamics in the Poisson manifold $\mathbb{L}^*$, with the expected energy $\langle H\rangle$ as the *classical* Hamiltonian. In particular, as can also be seen directly, the expected energy is conserved.

2.3 The classical approximation

> *Es ist wünschenswert, die folgende Frage möglichst elementar beantworten zu können: Welcher Rückblick ergibt sich vom Standpunkt der Quantenmechanik auf die Newtonschen Grundgleichungen der klassischen Mechanik?*
>
> Paul Ehrenfest 1927 [76, p. 455]

In the practice of quantum physics, one often approximates a complicated quantum dynamics by replacing certain quantities in the defining formulas by their expectations. This **classical approximation** is frequently justified and leads to more tractable simplified quantum systems. Making this replacement for all operators in the Ehrenfest equation leads to a corresponding classical dynamics.

As an example we consider an interacting multiparticle quantum system with mass matrix M, position operator q, and momentum operator p, both of dimension n, with dynamics given by the Hamiltonian $H = \frac{1}{2}p^T M^{-1}p + V(q)$. To arrive at an approximate classical equation of motion for the q-expectation $\overline{q} = \langle q\rangle$, we apply the Ehrenfest equation. Using the canonical commutation relations and componentwise expectations, we find the formulas

$$\frac{d}{dt}\langle q\rangle = M^{-1}\langle p\rangle, \quad \frac{d}{dt}\langle p\rangle = -\langle \nabla V(q)\rangle;$$

hence the equation

$$M\frac{d^2}{dt^2}\overline{q} + \langle \nabla V(q)\rangle = 0 \tag{2.15}$$

7 In terms of the symplectic Poisson algebra considered by Strocchi [280] to express the quantum mechanical dynamics of pure states ψ, the present Poisson algebra corresponds to the Poisson subalgebra of even functions of ψ.

by Ehrenfest [76], who observed the close formal relationship with the classical equation of motion

$$M\frac{d^2}{dt^2}q + \nabla V(q) = 0 \tag{2.16}$$

for this Hamiltonian. To turn this formal relationship into a quantitative approximation, we first prove the following bound in terms of the uncertainties (3.1):

Theorem 2.3.1 (Approximation lemma). *Let f be a twice continuously differentiable complex-valued function on $\mathbb{R}^n$. Then, for every vector q of n commuting self-adjoint quantities with convex joint spectrum and every state, we have (with the spectral norm)*

$$|f(q) - f(\overline{q})| \le \frac{1}{2}\|f''(q)\| \sum_{k=1}^{n} \sigma_{q_k}^2. \tag{2.17}$$

Proof. Indeed, for any $\widetilde{q}$ in the joint spectrum of q and $\varepsilon = \widetilde{q} - \overline{q}$, we have

$$f(\widetilde{q}) = f(\overline{q} + \varepsilon) = f(\overline{q}) + f'(\overline{q})\varepsilon + \int_0^1 \varepsilon^T f''(\overline{q} + s\varepsilon)\varepsilon s\, ds.$$

By assumption, $\overline{q} + s\varepsilon$ is for all $s \in [0,1]$ in the joint spectrum of q. Hence, by definition of the spectral norm

$$\left|\frac{\varepsilon^T f''(\overline{q} + s\varepsilon)\varepsilon}{\varepsilon^T}\varepsilon\right| \le \|f''(\overline{q} + s\varepsilon)\|_2 \le \|f''(q)\|.$$

Therefore,

$$\begin{aligned}|f(\widetilde{q}) - f(\overline{q}) - f'(\overline{q})(\widetilde{q} - \overline{q})| &\le \int_0^1 \|f''(q)\|\varepsilon^T \varepsilon s\, ds \\ &= \frac{1}{2}\|f''(q)\|\varepsilon^T\varepsilon = \frac{1}{2}\|f''(q)\| \sum_{k=1}^{n} (\widetilde{q}_k - \overline{q}_k)^2.\end{aligned}$$

This inequality therefore also holds for q in place of $\widetilde{q}$. Taking q-expectations, we find

$$\begin{aligned}|\langle f(q) - f(\overline{q})\rangle| &= |\langle f(q) - f(\overline{q}) - f'(\overline{q})(q - \overline{q})\rangle| \\ &\le \frac{1}{2}\|f''(q)\| \sum_{k=1}^{n} \langle (q_k - \overline{q}_k)^2\rangle = \frac{1}{2}\|f''(q)\| \sum_{k=1}^{n} \sigma_{q_k}^2.\end{aligned}$$

□

Returning to our original goal, we rewrite (2.15) in the form

$$M\frac{d^2}{dt^2}\overline{q} + \nabla V(\overline{q}) = -\langle \nabla V(q) - \nabla V(\overline{q})\rangle$$

and apply the approximation lemma to the right-hand side. Under the assumption that the potential V is three times continuously differentiable and the spectrum of the third derivative $V'''(q)$ is bounded by a constant C, we find the differential inequality

$$\left| M \frac{d^2}{dt^2} \overline{q} + \nabla V(\overline{q}) \right| \leq C \sum_{k=1}^{n} \sigma_{q_k}^2 .$$

Thus, as long as the uncertainties σ_{q_k} remain sufficiently small, the classical dynamical law (2.16) holds with good accuracy for the q-expectation $\overline{q}$ in place of q.

Under these conditions, which hold by the weak law of large numbers whenever q refers to the center of mass coordinates of macroscopic spherical bodies at macroscopic distances from each other, the q-expectations satisfy the traditional classical equation of motion.

This proves that Newton's mechanics is a macroscopic approximation to the quantum dynamics of q-expectations.

A similar analysis frequently allows certain quantities to be replaced by their expectations, leading to more tractable approximations. See, e.g., Ginibre [97] and Lieb & Seiringer [173].

2.4 The Rydberg–Ritz combination principle

Here we show that in any quantum system, the differences of the energy levels (the eigenvalues of the Hamiltonian H) are in principle directly observable, since they represent excitable oscillation frequencies of the system; thus can be probed by coupling the system to a harmonic oscillator with adjustable frequency. Therefore, the observed spectral properties of quantum systems appear as natural resonance phenomena.

To see this, we shall assume for simplicity a quantum system, whose Hamiltonian has a purely discrete spectrum. For a partially continuous spectrum, analogous results, in which sums are replaced by Stieltjes integrals, can be proved using the Gel'fand–Maurin theorem, also known under the name nuclear spectral theorem (see Maurin [182]).

We work in the Heisenberg picture on the basis of eigenstates of the Hamiltonian, such that $H|k\rangle = E_k|k\rangle$ for certain energy levels E_k. The q-expectation

$$\langle A(t) \rangle = \operatorname{Tr} \rho A(t) = \sum_{j,k} \rho_{jk} A_{kj}(t)$$

is a linear combination of the matrix elements

$$\begin{aligned} A_{kj}(t) &= \langle k|A(t)|j\rangle = \langle k|e^{iHt/\hbar} A e^{-iHt/\hbar}|j\rangle \\ &= e^{iE_k t/\hbar} \langle k|A|j\rangle e^{-iE_j t/\hbar} = e^{i\omega_{kj} t} \langle k|A|j\rangle, \end{aligned}$$

where

$$\omega_{kj} = \frac{E_k - E_j}{\hbar}. \tag{2.18}$$

Thus, the q-expectation exhibits quasiperiodic oscillatory behavior, whose frequencies ω_{jk} are scaled differences of energy levels. This relation, the modern form of the **Rydberg–Ritz combination principle** found in 1908 by RITZ [247], may be expressed in Planck's form[8]

$$\Delta E = \hbar\omega. \tag{2.19}$$

To probe the spectrum of a quantum system, we bring it into contact with a macroscopically observable (hence classically modeled) weakly damped harmonic oscillator. For simplicity, we treat just a single harmonic oscillator. In practice, one often observes many oscillators simultaneously, for example, by observing the oscillations of the electromagnetic field in the form of electromagnetic radiation—light, X-rays, or microwaves. However, in most cases, the oscillators may be regarded as independent and noninteracting. The result of probing a system with multiple oscillators results in a linear superposition of the results of probing with a single oscillator. This is a special case of the general fact that solutions of linear differential equations depend linearly on the right hand side.

From the point of view of the macroscopically observable classical oscillator, the probed quantum system appears simply as a time-dependent external force $F(t)$ that modifies the dynamics of the free harmonic oscillator. Instead of the harmonic equation $m\ddot{q} + c\dot{q} + kq = 0$ with real $m, c, k > 0$, we get the differential equation describing the **forced harmonic oscillator**, given by

$$m\ddot{q} + c\dot{q} + kq = F(t),$$

where the external force F is the q-expectation

$$F(t) = \langle A(t) \rangle$$

of a quantity A from the probed system. We assume the oscillator to have an adjustable frequency

$$\omega = \sqrt{\frac{k}{m}} > 0$$

and consider the response as a function of ω at fixed mass m and stiffness $k = m^2\omega$.

If the observation is done far from the probed system, such as an observation of light (electromagnetic radiation) emitted by a far away source (for example, a star, but

8 The formula (2.19) appears first in the famous 1900 paper by PLANCK [238] on the radiation spectrum of a black body. Planck wrote it in the form $\Delta E = h\nu$, where $h = 2\pi\hbar$, and $\nu = \omega/2\pi$ is the linear frequency. The symbol for the quotient $\hbar = h/2\pi$, which translates this into our formula was invented much later, in the 1930 quantum mechanics book by DIRAC [70].

also a Bunsen flame observed by the eye), the back reaction of the classical oscillator on the probed system can be neglected. Then the probed system can be considered as isolated and evolves according to the preceding analysis. Hence, the external force F can be written as a superposition

$$F(t) = \sum_l F_l e^{i\omega_l t}$$

of exponentials oscillating with the (positive and negative) Rydberg–Ritz frequencies, rearranged in linear order. The solution to the differential equation consists of a particular solution and a solution to the homogeneous equation. Due to damping, the latter is transient and decays to zero. There is a distinguished particular solution persisting after the transient decayed, which oscillates with the same frequencies as the force, easily seen to be given by

$$q(t) = \sum_l q_l e^{i\omega_l t}, \quad q_l = \frac{F_l}{m(\omega^2 - \omega_l^2) + ic\omega_l}.$$

Since the frequencies are real and distinct, the denominator cannot vanish. The energy in the lth mode is therefore proportional to the amplitude

$$|q_l|^2 = \frac{|F_l|^2}{m^2(\omega^2 k - \omega_l^2)^2 + c^2\omega_l^2}, \tag{2.20}$$

with a maximum at the resonance frequency $\omega = |\omega_l|$. The total energy is proportional to

$$\big|q(t)\big|^2 = \sum_l |q_l|^2 + \sum_{k\neq l} q_k^* q_l e^{i(\omega_k - \omega_l)t}. \tag{2.21}$$

We now look at the short-time average (recorded by a typical detector). If the frequencies ω_k with significant intensity are well-separated, the oscillating terms in (2.21) cancel out, and we find a total mean energy proportional to

$$a(t) \approx \sum_l |q_l|^2 = \sum_l \frac{|F_l|^2}{(m^2(\omega^2 k - \omega_l^2)^2 + c^2\omega_l^2}.$$

As a function of the varying frequency, this has the typical form of a spectral intensity, a superposition of Lorentz-shaped resonance curves, with local maxima very close to the resonance frequencies $|\omega_l|$.

2.5 The pure state idealization

A state ρ is called **pure** at time t if $\rho(t)$ has rank one, that is, maps the Hilbert space $\mathbb{H}$ to a 1-dimensional subspace, and **mixed** otherwise.

Although much of traditional quantum physics is phrased in terms of pure states, this is a very special case. In most actual experiments, the systems are open and the

states are mixed. Pure states are relevant only if they come from the ground state of a Hamiltonian, in which the first excited state has a large energy gap. Indeed, assume for simplicity that H has a discrete spectrum. In an orthonormal basis of eigenstates ϕ_k, functions $f(H)$ of the Hamiltonian H are defined by

$$f(H) = \sum_k f(E_k)\phi_k\phi_k^*$$

whenever the function f is defined on the spectrum. The equilibrium density is the canonical ensemble

$$\rho(T) = Z(T)^{-1}e^{-H/kT} = Z(T)^{-1}\sum_k e^{-E_k/kT}\phi_k\phi_k^*;$$

here k is the **Boltzmann constant**. (Of course, equating this ensemble with equilibrium in a closed system is an additional step beyond our system of axioms, which would require justification.) Since the trace equals 1, we find

$$Z(T) = \sum_k e^{-E_k/kT},$$

the textbook formula for the so-called **partition function**. In the limit $T \to 0$, all terms $e^{-E_k/kT}$ become 0 or 1, with 1 only for the k, corresponding to the states with least energy. Thus, if the ground state ϕ_1 is unique,

$$\lim_{T\to 0}\rho(T) = \phi_1\phi_1^*.$$

This implies that for low enough temperatures, the equilibrium state is approximately pure. The larger the gap between (nondegenerate) ground state and the first excited state, the better is the approximation at a given nonzero temperature. In particular, the approximation is good if the energy gap exceeds a small multiple of $E^* := kT$.

States of sufficiently simple systems (that is, those with a few energy levels only) can often be prepared in nearly pure states, by realizing a source governed by a Hamiltonian, in which the first excited state has a much larger energy than the ground state. Dissipation then brings the system into equilibrium, and as seen above, the resulting equilibrium state is nearly pure. Those low lying excited states, for which a selection rule suppresses the transition to a lower energy state, can be made nearly pure in the same way.

2.6 Schrödinger equation and formal Born rule

To see how the more traditional setting in terms of the Schrödinger equation arises, we consider the special case of a closed system in a pure state $\rho(t)$ at some time t. The **state vector** of such a system at time t is, by definition, a unit vector $\psi(t)$ in the range

of the pure state $\rho(t)$. It is determined up to a phase factor (of absolute value 1), and one easily verifies that

$$\rho(t) = \psi(t)\psi(t)^*. \tag{2.22}$$

Remarkably, under the dynamics for a closed system specified in the above axioms, this property persists with time if the system is closed and the state vector satisfies the **time-dependent Schrödinger equation**

$$i\hbar\dot{\psi}(t) = H\psi(t).$$

Thus, the state remains pure at all times. Conversely, for every pure state, the phases of $\psi(t)$ at all times t can be chosen such that the Schrödinger equation holds; the density operator is independent of this phase.

Moreover, if X is a vector of q-observables with commuting components and the spectrum of X is discrete, then the measure from Axiom (A5) is discrete:

$$\int d\mu(X) f(X) = \sum_k p_k f(X_k)$$

with spectral values X_k and nonnegative numbers p_k summing to 1, called **q-probabilities**.[9] Associated with the p_k are eigenspaces $\mathbb{H}_k$ such that

$$X\psi = X_k\psi \quad \text{for } \psi \in \mathbb{H}_k,$$

and $\mathbb{H}$ is the direct sum of the $\mathbb{H}_k$. Therefore, every state vector ψ can be uniquely decomposed into a sum

$$\psi = \sum_k \psi_k, \quad \psi_k \in \mathbb{H}_k.$$

ψ_k is called the **projection** of ψ to the eigenspace $\mathbb{H}_k$. If all eigenvalues of X are discrete and nondegenerate, each $\mathbb{H}_k$ is 1-dimensional and spanned by a normalized eigenvector ϕ_k. Then $X\phi_k = X_k\phi_k$, and the projection is given by $\psi_k = P_k\psi$ with the orthogonal projector $P_k := \phi_k\phi_k^*$, so that

$$\psi = \sum_k \phi_k\phi_k^*\psi.$$

A short calculation using Axiom (A5) now reveals that for a pure state (2.22), the q-probabilities p_k are given by the **formal Born rule**

$$p_k = |\psi_k(t)|^2 = |\phi_k^*\psi(t)|^2, \tag{2.23}$$

where $\psi_k(t)$ is the projection of $\psi(t)$ to the eigenspace $\mathbb{H}_k$.

At present, the formal Born rule (2.23) is just a piece of uninterpreted mathematics with suggestive naming. Questions of interpretation are discussed in Part II.

9 This leaves open the precise physical meaning of q-probabilities, and the question of how to measure them. This is the price to pay for not entering into interpretational issues at this stage.

3 Uncertainty, statistics, probability

In this chapter, we consider the way quantum theory represents deterministic, statistical, and probabilistic aspects of Nature. After a discussion in Section 3.1 of uncertainty in general, we discuss in Sections 3.2–3.4 several formal notions of classical probability, their relation to the probability concept used in applied statistics, and their dependence on the description used. Section 3.5 briefly discusses various examples, where probabilistic features emerge from fully deterministic situations—a theme to be taken up in much more detail in Chapter 11. Section 3.6 then shows how the statistical aspects of the quantum formalism naturally follow from the weak law of large numbers. The notions of c-probability in the classical case and of q-probability in the quantum case are formally defined in Section 3.7.

Historically, the concept of classical probability (including its use in stochastic processes) was given an undisputed formal mathematical foundation in 1933 in terms of the measure-theoretic setting of Kolmogorov [164]. Apart from this traditional axiomatic foundation of classical probability theory, there is a less well-known equivalent axiomatic treatment by Whittle [302] in terms of expectations. Here probabilities appear as the expectations of statements, $\{0,1\}$-valued random variables. von Plato [240] discusses the history of the concept of probability. Krüger et al. [169] discuss the history of probability in the various fields of application. Sklar [275] discusses the philosophical problems of the probability concept, with an emphasis on statistical mechanics.

3.1 Uncertainty

> *A quantity in the general sense is a property ascribed to phenomena, bodies, or substances that can be quantified for, or assigned to, a particular phenomenon, body, or substance. [...] The value of a physical quantity is the quantitative expression of a particular physical quantity as the product of a number and a unit, the number being its numerical value.*
>
> Guide for the Use of the International System of Units (Taylor [283])

> *The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories according to the way in which their numerical value is estimated.*
>
> *Type A. Those which are evaluated by statistical methods*
>
> *Type B. Those which are evaluated by other means*
>
> *[...] The quantities u_j^2 may be treated like variances and the quantities u_j like standard deviations.*
>
> NIST Reference on Constants, Units, and Uncertainty [220]

Uncertainty permeates all of human culture, not only science. Everything quantified by real numbers (as opposed to counting) is intrinsically uncertain because we cannot determine a real number with arbitrary accuracy. Even counting objects or events is uncertain in as much the criteria that determine the conditions under which some-

https://doi.org/10.1515/9783110667387-003

thing is counted are ambiguous. (When does the number of people in a room change by one while someone enters the door?)

We take the virtually universal presence of uncertainty as the most basic fact of science and give it a quantitative expression. Some of this uncertainty can be captured by probabilities and statistics, but the nature of much of this uncertainty is conceptual. Thus uncertainty is a far more basic phenomenon than statistics. It is an uncertainty in the notion of measurability itself. What does it mean to have measured something?

To be able to answer this we first need clarity in the terminology. To eliminate any trace of observer issues[1] from the terminology, we use the word **quantity** (as recommended in the above quote from the "Guide to the International System of Units") or – in a more technical context – **q-observable**[2] whenever quantum tradition uses the word observable. Similarly, to eliminate any trace of a priori statistics from the terminology, we frequently use the terminology **uncertain value** (in [214] simply called value) instead of q-expectation value, and **uncertainty** instead of q-standard deviation.

For the sake of definiteness, we first consider the notion of uncertain position. This may mean two things:

1. It may mean that the position could be certain, as in classical Newtonian physics, except that we do not know the precise value. However, observations of arbitrary accuracy are at least conceivable.
2. It may mean that the position belongs to an extended object, such as a neutron star, the Sun, a city, a cloud, a house, a tire, an apple, or a water wavelet. In each case, there is a clear approximate notion of position as a more or less fuzzy region in space, but it does not make sense to specify this position by coordinates to within millimeter accuracy.

It seems to be impossible to interpret the second case naturally in terms of the first case. The only physically distinguished point-like position of an extended object is its center of mass. Classically, one could therefore think of defining the exact position of an extended object to be the position of its center of mass. But the sun, a city, a house, or a water wavelet do not even have a well-defined boundary. Therefore, even

1 Except when relating to tradition, we deliberately avoid the notion of observables, since it is not clear on a fundamental level what it means to "observe" something, and since many things (such as the fine structure constant, neutrino masses, decay rates, scattering cross sections) observable in Nature are only indirectly related to what is traditionally called an "observable" in quantum physics.

2 Note that renaming notions has no observable consequences, but strongly affects the interpretation. To avoid confusion, this book follows the convention of Allahverdyan et al. [7] and adds the prefix "q-" to all traditional quantum notions that get here a new interpretation, and hence a new terminology. In particular, we use the terms q-observable, q-expectation, q-variance, q-standard deviation, q-probability, q-ensemble for the conventional terms observable, expectation, variance, standard deviation, probability, and ensemble.

the definition of their center of mass, which depends on what precisely belongs to the object, is ambiguous. And, is a tire really located at its center of mass—which is well outside the material the tire is made of? Things get worse in the microscopic realm, where the center of mass of a system of quantum particles has not even an exactly numerically definable meaning.

On closer inspection it seems that the situation of case 2 is very frequent in practice. Indeed, it is the *typical* situation in the macroscopic, classical world. Case 1 appears to be simply a convenient but unrealistic idealization.

Thus, uncertainty is only partially captured through statistical techniques. The latter apply only in case of highly repetitive uncertain situations, leading to a particular kind of uncertainty called aleatoric uncertainty (see, e. g., [67, 229]). More general kinds of uncertainty are discussed in the NIST Reference on Constants, Units, and Uncertainty [220], which may be regarded as the de facto scientific standard for representing uncertainty. This source explicitly distinguishes between uncertainties "which are evaluated by statistical methods" and those "which are evaluated by other means". For the second category, it is recognized that the uncertainties are not statistical, but should be treated "like standard deviations".

We capture this fundamental—not further explained but intuitive—notion of uncertainty at some time t in terms of q-expectations $\langle A\rangle_t$, and drop for simplicity the index t. The formulation chosen belongs to the formal core of quantum physics, as it is deliberately independent of the interpretation-sensitive notions of randomness, knowledge, observation, and measurement.

(GUP) General uncertainty principle: *A Hermitian*[3] *quantity A has the* ***uncertain value*** $\overline{A} = \langle A\rangle$ *with an* ***uncertainty*** *of*[4]

$$\sigma_A := \sqrt{\langle (A-\overline{A})^2\rangle} = \sqrt{\langle A^2\rangle - \overline{A}^2}. \tag{3.1}$$

In particular, the uncertain value $\overline{A}$ *is informative whenever its uncertainty* σ_A *is much less than* $|\overline{A}|$.

As discovered by Heisenberg [117], quantum physics predicts unavoidable uncertainty: The uncertainty of corresponding components of position q and momentum p cannot be both arbitrarily small since

$$\sigma_{p_j}\sigma_{q_j} \geq \frac{1}{2}\hbar.$$

3 The uncertain value and its uncertainty makes also sense in the nonhermitian case, but the uncertainty must be defined in this case as

$$\sigma_A := \sqrt{\left\langle (A-\overline{A})^*(A-\overline{A})\right\rangle} = \sqrt{\left\langle A^*A\right\rangle - \left|\overline{A}^2\right|}.$$

4 The equivalence of both expressions defining σ_A follows from $\overline{A} = \langle A\rangle$ and

$$\left\langle (A-\overline{A})^2\right\rangle = \left\langle A^2 - A\overline{A} - \overline{A}A + \overline{A}^2\right\rangle = \left\langle A^2\right\rangle - \langle A\rangle\overline{A} - \overline{A}\langle A\rangle + \overline{A}^2 = \left\langle A^2\right\rangle - \overline{A}^2.$$

This inequality is the famous **Heisenberg uncertainty relation—** not to be confused with the general uncertainty principle (GUP) defined above. The Heisenberg uncertainty relation is a special case of the more general statement (due to Robertson [249]) that for noncommuting Hermitian operators A, B,

$$\sigma_A \sigma_B \geq \frac{1}{2}|\langle [A,B]\rangle|, \tag{3.2}$$

which follows from the definitions.[5]

Like Ehrenfest's equation, the Heisenberg uncertainty relation does not involve interpretation-sensitive notions, hence belongs to the formal core of quantum physics. Both are valid independent of the interpretation of the q-expectations and their uncertainties.

Note that much of quantum physics can be developed without using the notion of probability at all. The notion of uncertain values $\langle A\rangle$ suffices for almost all of spectroscopy, quantum chemistry, and quantum statistical mechanics. For equilibrium statistical mechanics, this can be seen from the treatment in Neumaier & Westra [214].

3.2 Expectations as properties of anonymous events

Let x_k $(k = 1,\dots,N)$ denote the (real) values of some property of a collection of N similar classical objects. If the detailed identification of the objects is deemed irrelevant for certain purposes, the assignment of indices to the individual objects may be dropped, thereby anonymizing the data. Indeed, this is a common procedure in the statistical practice of handling sensitive data. Once this is done, we can no longer say which property belongs to which object—in the resulting description, the objects have become anonymous, or **indistinguishable**.

As a consequence, the individual values x_k play no longer a useful role in the anonymized collection. From a mathematical point of view, only symmetric functions of the x_k retain meaningful information about the collection. By a well-known theorem, every symmetric polynomial (and by taking limits, therefore any symmetric analytic function) of the x_k can be written as a function of the power sums $\sum x_k^e$ $(e = 1,2,3,\dots)$, equivalently, as a function of the sample expectations $\langle x^e\rangle = N^{-1}\sum_{k=1}^N x_k^e$. Some discontinuous symmetric functions also play a role, and can be written as a function of sample expectations of discontinuous functions. Thus, all **properties** of the

5 Indeed, the relation remains unchanged when subtracting from A and B its q-expectation, hence it suffices to prove it for the case, where both q-expectations vanish. In this case, $\langle A^2\rangle = \sigma_A^2$ and $\langle B^2\rangle = \sigma_B^2$, and the Cauchy–Schwarz inequality gives $|\langle AB\rangle|^2 \leq \langle A^2\rangle\langle B^2\rangle = \sigma_A^2\sigma_B^2$; hence $|\langle AB\rangle| \leq \sigma_A\sigma_B$. On the other hand, one easily checks that $i\,\mathrm{Im}\langle AB\rangle = \frac{1}{2}\langle [A,B]\rangle$, so that $\frac{1}{2}|\langle [A,B]\rangle| = |\,\mathrm{Im}\langle AB\rangle| \leq |\langle AB\rangle|$. Combining both inequalities gives the assertion.

anonymized collection are encoded in expectations $\langle f(x)\rangle$ of functions of the anonymous value x of an anonymous object of the collection.

It is precisely this situation that probability theory and statistics cater for—the description of anonymous events, not that of actual events! We assign probabilities to anonymous events such as "casting a die gives a six" (where the indefinite article indicates an anonymous die), not to the number of eyes shown on a particular die cast at a particular time (which is not a random variable but a fixed, though possibly unknown, value). We estimate the expected lifetime of "a 45 year old French male", not that of Francois Renon from Calais, say. And so on. Formally, from what is mathematically modeled, anonymous objects (whose only properties are expectation values and probabilities) are very different from **typical** objects, which are identifiable examples of particular objects (whose properties are individual values within observable typical ranges).

3.3 Classical probability via expectation

This section gives an elementary introduction to classical probability theory along the lines of WHITTLE [302], similar in spirit to the formal core of quantum mechanics.

Let Ω denote a finite or infinite set of labels of anonymous objects. We call the elements $\omega \in \Omega$ **experiments**.[6] Let $\mathbb{E}$ be a vector space of real valued functions on Ω containing the constant functions. We identify constant functions with their function values. The elements $A \in \mathbb{E}$ are called (real) **random variables**, and the value $A(\omega)$ is called the **realization** of A in experiment ω. The procedure that defines how to obtain the realization $f(\omega)$ for any experiment ω is called the **protocol** defining the random variable f. Functions, order relations, unary and binary operations, and limits on random variables are defined pointwise. Note, however, that a pointwise function $f(A)$ of a random variable $A \in \mathbb{E}$ does not necessarily lie in $\mathbb{E}$.

A **sample** is a finite set S of $|S| > 0$ experiments. The associated **sample mean** of a random variable A is defined by

$$\langle A\rangle_S := \frac{1}{|S|}\sum_{\omega\in S} A(\omega). \tag{3.3}$$

It is easily checked that any sample mean $\langle\cdot\rangle = \langle\cdot\rangle_S$ satisfies the following rules:

6 This makes the term "experiment" a formal object independent of its meaning in experimental physics, though the relationship is suggestive. In probability theory, Ω is called the **sample space**. In statistics, the $\omega \in \Omega$ may be identified with actual experiments carried out in the past or in the future. In classical statistical mechanics, the ω are only hypothetical experiments from a fictitious ensemble in the sense of Gibbs, as discussed in Section 15.3 of the Appendix. In general, Ω is just a set, and calling the $\omega \in \Omega$ experiments is just a convenient intuition without formal meaning (comparable to calling elements of a vector space vectors, even though they may be functions or matrices). For examples, see Section 3.5 below.

(E1) $\langle 1\rangle = 1$.
(E2) $\langle \alpha A + \beta B\rangle = \alpha\langle A\rangle + \beta\langle B\rangle$ for $\alpha, \beta \in \mathbb{R}$.
(E3) $A \geq 0$ implies $\langle A\rangle \geq 0$.
(E4) $A \geq 0$, $\langle A\rangle = 0$ implies $A = 0$.
(E5) $A_k \downarrow 0$ implies $\langle A_k\rangle \downarrow 0$.

Here $\downarrow$ denotes pointwise convergence from above. To abstract from a particular sample, we define a **stochastic model** as an arbitrary mapping that assigns to each random variable A a real number $\langle A\rangle$, called the **expectation** (or **expected value** or **mean**) of A, such that the axioms (E1)–(E5) hold whenever the expectations in question exist. All samples, arbitrary convex combinations of samples, and their limits, define a stochastic model. Given real statistical data from real experiments, the quality of a stochastic model is assessed by how well the expectations of key random variables match corresponding sample expectations for samples drawn at random in some informal sense.

As a simple consequence of the axioms, we note:
(E6) $\langle A^2\rangle = 0$ implies $A = 0$.
(E7) $A \leq B$ implies $\langle A\rangle \leq \langle B\rangle$.

Indeed, (E6) follows directly from (E4). For (E7), the assumption gives $B - A \geq 0$. Hence $\langle B - A\rangle \geq 0$ by (E3); hence $\langle B\rangle - \langle A\rangle \geq 0$ by (E2), giving $\langle A\rangle \leq \langle B\rangle$.

An example of a random variable is the number n of eyes on the top side of a die. Here $n(\omega) \in \{1, \dots, 6\}$ is the number of eyes on the die visible in experiment ω. We may thus consider $\mathbb{E}$ to be the algebra generated by a single random variable $n = n(\omega)$, taking the values $1, 2, 3, 4, 5, 6$. Thus, the relevant random variables are the functions $A = A(n)$, defined by

$$A(n)(\omega) := A\big(n(\omega)\big).$$

A is determined by the vector of the six values $A_1 = A(1), \dots, A_6 = A(6)$. Therefore, we may identify $\mathbb{E}$ with the vector space $\mathbb{R}^6$ with componentwise operations. The stochastic model, defined by

$$\langle A\rangle := \frac{1}{6}(A_1 + \cdots + A_6),$$

models an ideal, permutation symmetric die.

In practice, the set Ω may be different depending on the imagination of people and the intended use. The expectation value is independent of $\omega \in \Omega$ and depends—as in the example just given—on the vector space $\mathbb{E}$ of relevant random variables only. This algebra is always commutative and associative.

A **statement** is a $\{0, 1\}$-valued random variable A. The statement is **true** (**false**) in an experiment ω if $A(\omega) = 1$ (that is, $A(\omega) = 0$). The **probability** of a statement A,

defined as

$$\Pr(A) := \langle A \rangle,$$

is a number between 0 and 1. Indeed, we have $0 \le A \le 1$. Thus, by (E3), $\langle A \rangle \ge 0$. By (E6) and (E1), $\langle A \rangle \le \langle 1 \rangle = 1$. An example for a statement is $S = [|A| \ge \alpha]$, for any random variable X and real number $\alpha \ge 0$. Here $[\ldots]$ denotes the statement defined by the formula inside the square brackets. Since $[|A| \ge \alpha] \le A^2/\alpha^2$, we conclude from (E7) the **Chebyshev inequality**

$$\Pr(|A| \ge \alpha) \le \langle A^2 \rangle / \alpha^2. \tag{3.4}$$

Proposition 3.3.1. *If $A_1, \ldots, A_n$ are alternative statements, of which exactly one is true in each experiment, then the probabilities $p_i := \Pr(A_i)$ sum up to 1, and $\Pr(A_i \wedge A_j) = 0$ for $i \ne j$.*

Proof. Indeed, the random variable $E := \sum_{i=1}^n A_i$ satisfies

$$\chi(\omega) = \sum_{i=1}^n A_i(\omega) = 1 \quad \text{for all } \omega \in \Omega,$$

because by definition exactly one A_i occurs in each experiment ω. Therefore, $E = 1$, and

$$1 = \langle 1 \rangle = \langle E \rangle = \left\langle \sum_{i=1}^n A_i \right\rangle = \sum_{i=1}^n \langle A_i \rangle = \sum_{i=1}^n p_i.$$

Similarly, $(A_i \wedge A_j)(\omega) = A_i(\omega) \wedge A_j(\omega) = 0$ for $i \ne j$, since at most one of $A_i(\omega)$ and $A_j(\omega)$ can be true. Hence, $A_i \wedge A_j = 0$ and $\Pr(A_i \wedge A_j) = \Pr(0) = \langle 0 \rangle = 0$. □

The **cumulative distribution function** (**CDF**) of a random variable A is the function $\mathrm{cdf} : \mathbb{R} \to [0,1]$, defined as

$$\mathrm{cdf}(x) := \Pr(A \le x) = \big\langle [A \le x] \big\rangle$$

for all x.

Proposition 3.3.2. *Every CDF is monotone increasing and satisfies*

$$\lim_{\varepsilon \downarrow 0} \mathrm{cdf}(x + \varepsilon) = \mathrm{cdf}(x), \quad \lim_{x \to -\infty} \mathrm{cdf}(x) = 0, \quad \lim_{x \to +\infty} \mathrm{cdf}(x) = 1. \tag{3.5}$$

Therefore, every CDF is continuous from the right.

Proof. By (E6), $x \le x'$ implies

$$[A \le x'] - [A \le x] \ge 0.$$

So by (E6),

$$\begin{aligned}\Pr(A \le x') - \Pr(A \le x) &= \langle [A \le x'] - [A \le x] \rangle = \langle [x < A \le x'] \rangle \\ &= \Pr([x < A \le x']) \ge 0,\end{aligned}$$

and thus

$$\operatorname{cdf}(x) \le \operatorname{cdf}(x').$$

To prove continuity from the right, we note that the random variable

$$B(\varepsilon) := [x < A \le x + \varepsilon]$$

vanishes at ω with $A(\omega) \le x$ and for $\varepsilon < A(\omega) - x$ if $A(\omega) > x$, and is 1 otherwise. Therefore, $B(\varepsilon) \downarrow 0$ for $\varepsilon \downarrow 0$. Thus,

$$\langle B(\varepsilon) \rangle \downarrow 0 \quad \text{for } \varepsilon \downarrow 0$$

by (E5). But

$$\langle B(\varepsilon) \rangle = \Pr(x < A \le x + \varepsilon) = \operatorname{cdf}(x + \varepsilon) - \operatorname{cdf}(x).$$

Hence, the first limit of (3.5) follows. The other limits are proved in the same way. □

Let A be a random variable with cumulative distribution function cdf. If f is a step function with finitely many jump points at $x_1 \le \cdots \le x_n$, and f is continuous from the left, then

$$\langle f(A) \rangle = f(x_1)\operatorname{cdf}(x_1) + \sum_{k=1}^{n-1} f(x_{k+1})(\operatorname{cdf}(x_{k+1}) - \operatorname{cdf}(x_k)). \tag{3.6}$$

Indeed, the right hand side equals the expectation of

$$f(x_1)[A \le x_1] + \cdots + \sum f(x_k)[x_{k-1} < A \le x_k] = f(A).$$

By taking a continuum limit in (3.6), we obtain the **Stieltjes integral representation**

$$\langle f(A) \rangle = \int_{-\infty}^{\infty} f(\xi)\, d\operatorname{cdf}(\xi)$$

for the expectation of $f(A)$, for any function f that is continuous from the left. Thus, the CDF contains all information about expectations of functions of a real random variable.

For a random variable A with continuously differentiable cumulative distribution function, the **distribution** or **density** ρ of A is defined as

$$\rho(\xi) = \frac{d}{d\xi} \operatorname{cdf}(\xi).$$

It is always nonnegative since the CDF is monotone increasing. The CDF can be expressed in terms of the density as

$$\mathrm{cdf}(\xi) = \int_{-\infty.}^{\xi} \rho(\zeta)\, d\zeta$$

If the density $\rho(x)$ exists, it also carries all information about expectations of functions of A. Indeed, we have

$$\langle f(A) \rangle = \int_{-\infty}^{\infty} f(\xi)\rho(\xi)\, d\xi$$

since $d\,\mathrm{cdf}(\xi) = \rho(\xi)d\xi$.

3.4 Description dependence of probabilities

Classical probabilities are dependent on the description used. The latter encode the assumed knowledge about the system under study. Note that the implied concept of **knowledge** is not the knowledge of a particular observer or person, but an informal shorthand for what is modeled in a particular description. Thus, the "knowledge available", that is, the knowledge encoded into a particular description, is an objective property of the description used to model the system, independent of who "knows" or "uses" it. Nothing mental is implied. A change of knowledge is therefore just a change of the model used to describe a particular system.

Sample expectations, and hence sample probabilities, have a clear operational meaning. But they are properties of the specific sample taken—changing the sample changes the properties. For example, almost always when the values for new realizations become known, the values of sample expectations and sample probabilities change.

Cumulative distribution functions are easy to estimate on a sample S by the sample CDF

$$\mathrm{cdf}_S(\xi) = \frac{\text{number of } \omega \in S \text{ with } x(\omega) \le \xi}{|S|},$$

corresponding to the sample expectation (3.3). A sample CDF is always a step function with discontinuities at the $x(\omega)$ with $\omega \in S$. In many cases, it is well approximated by a smoothed CDF, using one of many smoothing methods available. The derivative of the smoothed CDF then serves as an estimate for the density of a random variable A in a stochastic model, in which A is treated as a random variable with density.

In the univariate case treated above in detail, going from a known sample to predictions for not yet known samples is based on the possibility to approximate sample distributions of different sample sizes by a single model distribution since their CDFs are very similar. This also holds in the multivariate case, though different techniques must be used to establish corresponding results. Finite sample properties can

be proved using arguments basically similar to our proof of the weak law of large numbers (3.11). Under reasonable assumptions, the uncertainty of the most relevant random variables scales with $O(N^{-1/2})$. This is the contents of large N approximations that take the form of laws of large numbers and central limit theorems. They guarantee enhanced predictability of the kind that the mean uncertainty is approximately the single case uncertainty divided by $\sqrt{N}$.

Thus, the precise meaning of expectations and probabilities depends on which stochastic model is used for a given situation. To see what happens when we change the description, we consider the concept of conditional expectation, which models the reweighing of evidence leading to a change of the description of a model. A **weight** is a nonzero random variable $P \geq 0$. Let

$$\Omega_P = \{\omega \in \Omega \mid P(\omega) > 0\}.$$

We project the algebra $\mathbb{E}$ of random variables $X : \Omega \to \mathbb{R}$ to the algebra $\mathbb{E}_P$ of random variables $Y : \Omega_P \to \mathbb{R}$ by means of the homomorphism

$$\cdot|_{\Omega_P} : \mathbb{E} \to \mathbb{E}_P : X \mapsto X|_{\Omega_P}$$

that restricts all random variables $X \in \mathbb{E}$ to $X|_{\Omega_P} \in \Omega_P$. The **conditional expectation** of $X : \Omega \to \mathbb{R}$ with respect to the weight P is defined as

$$\langle X|_{\Omega_P}\rangle_P = \langle X\rangle_P := \frac{\langle XP\rangle}{\langle P\rangle}.$$

Note that $\langle P\rangle > 0$ by (E4). The resulting mapping $\langle\cdot\rangle_P : \mathbb{E}_P \to \mathbb{R}$ satisfies all expectation axioms: (E1)–(E3) and (E5) follow directly from the corresponding axioms for $\langle\cdot\rangle$. Only (E4) is nontrivial; $\langle X^2\rangle_P = 0$ implies $\langle X^2P\rangle = \langle P\rangle\langle X^2\rangle_P = 0$. But $X^2P \geq 0$. Hence, (E4) gives $X^2P = 0$, that is, $X(\omega)^2P(\omega) = 0$ for all ω. Thus, $X(\omega) = 0$ whenever $P(\omega) = 1$. This gives $X|_{\Omega_P} = 0$ and proves (E4). Therefore $\langle\cdot\rangle_P$ is a proper expectation for the experiments in Ω_P.

In particular, any statement A with positive probability $\Pr(A) > 0$ may be considered as a weight. Since Ω_A is the set of all those experiments, where A is true, conditional expectation with respect to A is just expectation in the light of the assumed evidence that A is true. In particular, the probability for A being true changed to 1. More generally, the **conditional probability** of a statement B, given the statement A, is defined by

$$\Pr(B|A) := \langle B\rangle_A = \frac{\langle BA\rangle}{\langle A\rangle} = \frac{\Pr(A \wedge B)}{\Pr(A)}.$$

The last equality holds since

$$BA(\omega) = B(\omega)A(\omega) = \begin{cases} 1 & \text{if } A(\omega) \wedge B(\omega), \\ 0 & \text{otherwise.} \end{cases}$$

It is easily checked that for any statements A and B, we have

$$\Pr(B \wedge A) = \Pr(B|A)\Pr(A), \tag{3.7}$$

$$\Pr(A|B) = \Pr(A)\frac{\Pr(B|A)}{\Pr(B)}. \tag{3.8}$$

(3.8) is often called the **Bayes theorem**. In this context, $\Pr(A)$ is called the **prior probability** of A, $\Pr(A|B)$ its **posterior probability**, and $\Pr(B|A)/\Pr(B)$ the **update ratio**. If A and B are independent, then $\langle AB\rangle = \langle A\rangle\langle B\rangle$; hence $\Pr(B|A)/\Pr(B) = \langle AB\rangle/\langle A\rangle\langle B\rangle = 1$. Thus, the update ratio captures the degree to which the knowledge of B affects knowledge of A. Thus, Bayes theorem allows us to describe the change of probability of a class of statements when new information (namely the statement B) arrives and is accepted as valid. Bayes theorem is important to understand what it means to get new information when A is a statement of interest and B is information that becomes known. Bayes theorem tells how the probability of A changes under new insight. $\Pr(A)$ changes into $\Pr(A|B)$, so the probability of A must be multiplied by the update ratio $\Pr(B|A)/\Pr(B)$.

3.5 The stochastic description of a deterministic system

A stochastic description of a deterministic system is a reduced deterministic description by moments rather than details. Formally, it is obtained by restricting an algebra of commuting, classical quantities describing a given deterministic system to the subalgebra of quantities completely symmetric in some properties declared **indistinguishable** for the purposes of the reduced description. Such reduced (or coarse-grained) descriptions will be discussed in more detail in Chapter 7. In the simplest case, where expectation is defined as sample expectation, the individual realizations, over which the sample mean is taken are declared indistinguishable, with the consequence that only symmetric functions of realizations, hence functions of expectations, are available in the reduced descriptions.

In the terminology of knowledge discussed in Section 3.4, such a reduction amounts to forgetting or ignoring information known from the more detailed model. We have additional modeling uncertainty due to the lack of detail in the description used. This description is independent of a probabilistic interpretation.[7] It means that one only considers a limited family of well-behaved relevant quantities in place of the multitude of quantities in a more detailed description.

The analysis presented here allows one to apply statistical models to complicated *deterministic* situations—not only in physics—and to *single* complicated spatial events or time series. In each case, a suitable concept of expectation is introduced that allows one to make probabilistic and other statistical statements about deterministic situations.

7 Though via the construction of Section 3.7 below, it can always be given one in terms of c-probabilities.

Important examples of statistical models for deterministic situations with increasingly random appearance are:

(i) The deterministic but irregular sequence of prime numbers (Tennenbaum [285]). Here experiments are the natural numbers, and expectations are introduced through a mathematically rigorous limit.

(ii) Rounding errors in deterministic floating-point computations (Vignes [289]). Here experiments are sequences of floating-point operations, and expectations are introduced through an empirical model for single rounding errors and an assumption of independence.

(iii) Texture in a single (hence fully determined) picture (Heeger & Bergen [115]). Here experiments are hypothetical images of the same size, and expectations are introduced through a mean over a (not well-defined) neighborhood of pixels.

(iv) Economic time series, for example, prices of oil and electricity (Granger & Newbold [106]). Here experiments are hypothetical scenarios for the time series, and expectations are introduced through a no longer fully well-defined time average.

(v) Tomorrow's weather prediction, when based on classical fluid mechanics (Gneiting & Raftery [100]). Here experiments are again hypothetical weather scenarios, and expectations are introduced through an even less well-defined space-time average.

(vi) Deciding for red or black by spinning a roulette wheel (Hopf [138]). Here expectations may be introduced through symmetry arguments (for an ideal wheel), or through ergodic theory.

(vii) The classical statistical mechanics of an ideal gas. Here Boltzmann [41] introduced expectations through an average over all particles.

(viii) The classical statistical mechanics of solids and fluids. Here (see the discussion in Section 15.3 of the Appendix) Gibbs [96] introduced expectations through a fictitious average over many systems with identical macroscopic properties.

3.6 Deterministic and stochastic aspects of q-expectations

A quantity A is considered to be **significant** if $\sigma_A \ll |\overline{A}|$, while it is considered as **noise** if $\sigma_A \gg |\overline{A}|$. If A is a quantity and $\widetilde{A}$ is a good numerical approximation of its value, then $\Delta A := A - \widetilde{A}$ is noise. Sufficiently significant quantities can be treated as deterministic; the analysis of noise is the subject of statistics.

Statistics is based on the idea of obtaining information about noisy quantities of a system by repeated **sampling** from a **population**[8] of independent systems with iden-

8 Physicists usually speak of an ensemble in place of a population. In this book, the statistical term **population** is used instead, to keep the discussion unambiguous, since in connection with the micro-canonical, canonical, or grand canonical ensemble, the term **ensemble** is essentially synonymous with a density operator of a particularly simple form. The traditional notion of an ensemble is criticized in Section 15.3 of the Appendix.

tical preparation, but differing in noisy details not controllable by the preparation. In the present context, such systems are described by the same Hilbert space, the same set of quantities to be sampled, and the same state $\langle\cdot\rangle_0$. The quantities therefore belong to the algebra $\mathrm{Lin}\,\mathbb{H}_S$ of linear operators on a Euclidean space $\mathbb{H}_S$ dense in the Hilbert space of the system.

More precisely, stochastic features emerge when we consider a large sample of similar subsystems of a quantum system. For an ensemble of independent measurements on identically prepared systems, the consensus of all interpretations is that q-expectations represent (within the accuracy allowed by the law of large numbers) a statistical average of the measurement results.[9]

We now show that a q-expectation, though introduced as an intrinsic measure of uncertainty, may be viewed as a statistical property of many independent identically prepared systems. We regard the systems of the population considered as subsystems of a bigger system (for example, the laboratory) whose set of quantities is given by the algebra $\mathrm{Lin}\,\mathbb{H}$ of linear operators on a big Euclidean space $\mathbb{H}$ dense in the Hilbert space of the big system. To model identically prepared subsystems, we consider injective homomorphisms from $\mathrm{Lin}\,\mathbb{H}_S$ into $\mathrm{Lin}\,\mathbb{H}$ mapping, each reference quantity $A \in \mathrm{Lin}\,\mathbb{H}_S$ to the quantity $A_l \in \mathrm{Lin}\,\mathbb{H}$ of the lth subsystem considered to be "identical" with A. Of course, in terms of the big system, the A_l are not really identical; they refer to quantities distinguished by position and/or time. That the subsystems are **identically prepared** is instead modeled by the assumption

$$\langle A_k\rangle = \langle A_l\rangle \quad \text{for all } k \neq l, \tag{3.9}$$

and that they are **independent** by the assumption

$$\langle A_k A_l\rangle = \langle A_k\rangle\langle A_l\rangle \quad \text{for all } k \neq l. \tag{3.10}$$

The following result is fundamental for statistical considerations:

Theorem 3.6.1 (Weak law of large numbers). *For a sample of quantities A_l ($l = 1, \ldots, N$), satisfying* (3.9) *and* (3.10), *the **mean** quantity*

$$\widehat{A} := \frac{1}{N}\sum_{l=1}^{N} A_l$$

9 In order to take q-expectations alternatively as a time average of a single system, one would need to invoke an **ergodic theorem**, stating that the time average equals the ensemble average. However, most deterministic systems are far from ergodic. (This is mentioned, for example, in the statistical physics book by Landau & Lifschitz [172, Footnote 2, p. 12].) From the eight examples of statistical models for deterministic systems given in Section 3.5, only two—cases (vi) and (vii)—have a generally valid ensemble interpretation in terms of ergodicity. Thus, the interpretation of q-expectations as a time average is usually not warranted. Moreover, in the quantum version to be discussed in a moment, such an ergodic theorem makes sense only semiclassically. This also means (Neumaier [199]) that—in contrast to what is usually done in the popular literature—the so-called **vacuum fluctuations** of quantum field theory cannot be interpreted as fluctuations in time.

(which again is a quantity) satisfies for any l

$$\langle \widehat{A} \rangle = \langle A_l \rangle, \quad \sigma_{\widehat{A}} = \sigma_{A_l}/\sqrt{N}. \tag{3.11}$$

Proof. By (3.9) and (3.10), $\mu := \langle A_l \rangle$ and $\sigma := \sigma_{A_l}$ are independent of l, and we have

$$\begin{aligned} \langle \widehat{A} \rangle &= \frac{1}{N}(\langle A_1 \rangle + \cdots + \langle A_N \rangle) = \frac{1}{N}(\mu + \cdots + \mu) = \mu, \\ \langle \widehat{A}^* \widehat{A} \rangle &= \frac{1}{N^2} \left\langle \left(\sum_j A_j \right)^* \left(\sum_k A_k \right) \right\rangle = N^{-2} \sum_{j,k} \langle A_j^* A_k \rangle. \end{aligned} \tag{3.12}$$

Now

$$\langle A_j^* A_j \rangle = \langle A_j \rangle^* \langle A_j \rangle + \sigma_{A_j}^2 = |\mu|^2 + \sigma^2,$$

and by (3.10) for $j \neq k$,

$$\langle A_j^* A_k + A_k^* A_j \rangle = 2 \operatorname{Re} \langle A_j^* A_k \rangle = 2 \operatorname{Re} \langle A_j \rangle^* \langle A_k \rangle = 2 \operatorname{Re} \mu^* \mu = 2|\mu|^2.$$

The sum in (3.12) leads to a contribution of $|\mu|^2 + \sigma^2$ for each of the N diagonal elements, and of $2|\mu|^2$ for each of the $\binom{N}{2}$ pairs of off-diagonal elements. Therefore,

$$\langle \widehat{A}^* \widehat{A} \rangle = N^{-2} \left(N(|\mu|^2 + \sigma^2) + \binom{N}{2} 2|\mu|^2 \right) = N^{-1}\sigma^2 + |\mu|^2,$$

so that

$$\sigma_{\widehat{A}}^2 = \langle \widehat{A}^* \widehat{A} \rangle - \langle \widehat{A} \rangle^* \langle \widehat{A} \rangle = N^{-1}\sigma^2,$$

and the assertions follow. □

A significant body of work in probability theory shows that the conditions under which $\sigma_{\widehat{A}} \to 0$ as $N \to \infty$ can be significantly relaxed. Thus, in practice, it is sufficient if (3.9) and (3.10) are approximately valid.

The significance of the weak law of large numbers lies in the fact that (3.11) becomes arbitrarily small as N becomes sufficiently large. Therefore, the uncertainty of quantities—when averaged over a large population of identically prepared systems—becomes arbitrarily small, while the mean value reproduces the value of each quantity.

The weak law of large numbers implies that, in a context where many repeated experiments are feasible, states can be given a **frequentist** interpretation, in which $\overline{A} = \langle A \rangle$ is the **expectation** of A, empirically defined as an average over many realizations. In this case (and only in this case), the uncertainty σ_A becomes the standard deviation of A; then it captures the absolute accuracy of the individual realizations.

This determines the conditions under which deterministic and statistical reasoning are justified:

(SP) Statistical principle: *Deterministic reasoning is appropriate for all sufficiently significant quantities. Statistical reasoning is necessary for noisy quantities, and requires that these quantities are sufficiently similar and sufficiently independent to ensure that their mean is significant.*

3.7 What is probability?

In Section 3.3, we derived from (E1)–(E5) the traditional probabilistic machinery for single real random variables. More generally, it can be proved (see WHITTLE [302]) that even in the multivariate case, the above approach to classical probability is equivalent to the measure theoretical approach in the traditional axiomatic setting of Kolmogorov. In this equivalence, Ω is an abstract measure space; a stochastic model is a probability measure on Ω, and $\mathbb{E}$ is a vector space of random variables with finite expectations.

The exposition in WHITTLE [302] (or, in more abstract terms, already in GELFAND & NAIMARK [94]) shows that, if the X_j are pairwise commuting, it is possible to define for any Gibbs state in the present sense, random variables X_j in Kolmogorov's sense, such that the expectation of all sufficiently regular functions $f(X)$ defined on the joint spectrum of X agrees with the value of f. It follows that in the pairwise commuting case, it is always possible to construct a probability interpretation for the quantities, completely independent of any assumed microscopic reality. (If the components of X do not commute, a probabilistic interpretation in the Kolmogorov sense is no longer possible because of the nonclassical uncertainty relations (3.2).)

The details (which the reader unfamiliar with measure theory may simply skip) are as follows: We may associate with every vector X of quantities with commuting components a time-dependent, monotone linear functional $\langle\cdot\rangle_t$ defining the **expectation**

$$\langle f(X)\rangle_t := \operatorname{Tr}\rho(t)f(X)$$

at time t of arbitrary bounded continuous functions f of X. These functions define a commutative C^*-algebra $\mathbb{E}(X)$. The **spectrum** $\operatorname{Spec} X$ of X is the set of all $*$-homomorphisms (called **characters**) from $\mathbb{E}(X)$ to $\mathbb{C}$, and has the structure of a Hausdorff space, with the **weak-$*$ topology**, obtained by calling a subset S of $\operatorname{Spec} X$ closed if, for any pointwise convergent sequence (or net) contained in S, its limit is also in S. Now an expectation functional, satisfying (E1)–(E5), turns out to be equivalent to a multivariate probability measure $d\mu_t(X)$ (on the sigma algebra of Borel subsets of the spectrum Ω of X), defined by

$$\int d\mu_t(X)f(X) := \int \rho(t)f(X) = \langle f(X)\rangle_t.$$

Both Whittle's and Kolmogorov's foundations of classical probability theory are axiomatic, hence independent of the interpretation of the axioms. We may refer to probability as defined by Kolmogorov or Whittle as **c-probability**.

In generalization of this, we refer, for any Hermitian operators P satisfying $0 \le P \le 1$, to its expectation as the **q-probability**

$$\Pr(P) := \langle P\rangle$$

of P. As a special case, we call a quantity P, satisfying $P^2 = P = P^*$, a **statement**; then $0 \leq P \leq 1$ follows by the spectral theorem. For a statement P, the uncertainty of its probability $p = \overline{P}$ is $\sigma_P = \sqrt{p(1-p)}$ since by (3.1), $\sigma_P^2 = \langle P^2 \rangle - \overline{P}^2 = p - p^2$. Another special case is the q-probability of a self-adjoint Hermitian q-observable A taking values in some open interval $]a, b[$ of real numbers, defined as

$$\Pr(A \in]a, b[) := \langle P_{]a,b[}(A) \rangle,$$

where $P = P_{]a,b[}(A)$ is the spectral projector of A to the interval $]a, b[$. Note that here $P^2 = P = P^*$, so that P is a statement—the formal equivalent of the informal statement "A is in $]a, b[$".

Whittle's approach is essentially equivalent to the commutative case of the formal core of quantum mechanics, interpreted in statistical terms. Then q-probabilities and c-probabilities agree.

We discuss in more detail the important special case of binary tests, where Born's rule frequently applies essentially exactly. An **ideal binary measurement**, for example, the click of a detector, is described by a statement P, coding the presence (1) or absence (0) of a click. In particular, a **test for a state**[10] ϕ with $\phi^*\phi = 1$ is an ideal binary measurement of $P = \phi\phi^*$; it is easily checked that this is a statement. By the above, such a test turns out positive with probability $p = \langle P \rangle$. In particular, if the system is in a pure state ψ, then $p = \langle P \rangle = \psi^* P \psi = \psi^* \phi\phi^* \psi = |\phi^* \psi|^2$; hence

$$p = |\phi^* \psi|^2.$$

This is the well-known **squared probability amplitude** formula appearing in the formal Born's rule, which appears not as a basic axiom, but as one of its natural consequences.

10 Note that a test for ϕ turns out positively with probability 1 if the measured system is in the pure state ϕ. However, it also turns out positively with a positive probability if the measured system is in a pure state different from ϕ, as long as it is not orthogonal to it. Thus, calling it a "test for ϕ", though conventional, is a misnomer.

4 Euclidean spaces

There is a notational discrepancy in how mathematicians and physicists treat Hilbert spaces. In physics, one often works with finite-dimensional Hilbert spaces treated as $\mathbb{C}^n$, and hence wants to write the Hermitian inner product as $\langle x, y\rangle = x^* y$. This definition dictates the use of a Hermitian inner product that is antilinear in the first argument, the convention followed, for example, in REED & SIMON [246]. This is also the choice adopted in Dirac's bracket notation, whose usage in quantum mechanics is very widespread.

Our notation is chosen to extend—as closely as possible—the traditional notation of standard finite-dimensional matrix algebra to arbitrary complex inner product spaces and associated linear operators. In matrix algebra, column vectors and the corresponding matrices with one column are identical objects, row vectors are the linear functionals, and the adjoint is the conjugate transpose. For example, $\mathbb{H} = \mathbb{H}^\times = \mathbb{C}^n$ is the space of column vectors of size n, the dual space $\mathbb{H}^*$ is the space of row vectors of size n, and the operator product $\phi^*\psi$ of a row vector ϕ^* and a column vector ψ is the standard Hermitian inner product of the column vectors ϕ and ψ. We use Greek lower case letters to write vectors, thus emphasizing their intended use as quantum state vectors in quantum mechanics.

On the other hand, mathematicians working on reproducing kernel Hilbert spaces use an inner product (x, y) antilinear in the second argument, related to the physicist's inner product $\langle x, y\rangle$ by $(x, y) = \overline{\langle x, y\rangle}$. This is the convention followed, for example, in RUDIN [252]. Although the two ways of defining the inner product lead to fully equivalent theories, all details look a bit different, a fact that has to be taken into account when reading the literature on the subject. For example, in the description based on the physical tradition, it is preferable to work with the antidual space in place of the dual space used in the mathematical tradition.

In functional analysis, linear operators in Hilbert spaces are usually considered each with their own domain. But many computations in quantum mechanics require the consideration of algebras of operators with a common domain. The latter is a Euclidean space, a dense subspace of a Hilbert space. This space and its antidual play in many respects a more basic role in quantum physics than the Hilbert space itself.

Therefore, and to avoid possible confusion caused by the different traditions, we give in the present chapter a self-contained introduction to Euclidean spaces and their associated spaces. All proofs are carried out in detail.

4.1 Euclidean spaces and their antidual

A **Euclidean space** is a complex vector space $\mathbb{H}$ with a binary operation that assigns to $\phi, \psi \in \mathbb{H}$ the **Hermitian inner product** $\langle \phi, \psi\rangle \in \mathbb{C}$, antilinear in the first and linear

https://doi.org/10.1515/9783110667387-004

in the second argument, such that

$$\overline{\langle \phi, \psi \rangle} = \langle \psi, \phi \rangle, \tag{4.1}$$

$$\langle \psi, \psi \rangle > 0 \quad \text{for all } \psi \in \mathbb{H} \setminus \{0\}. \tag{4.2}$$

Here $\alpha > 0$ says that the complex number α is real and positive.

Since every Euclidean space can be completed to a Hilbert space (see Theorem 4.2.2), the Euclidean spaces are, in fact, just the subspaces of Hilbert spaces, with the induced inner product. However, it is of interest to develop the theory of Euclidean spaces independently since some additional topological structure is present that has no simple counterpart in the Hilbert space setting.

We define the **antidual** $\mathbb{H}^\times$ of $\mathbb{H}$ to be the vector space of antilinear functionals $\phi : \mathbb{H} \to \mathbb{C}$. We turn $\mathbb{H}^\times$ into a locally convex space (see RUDIN [252, Chapter 3]) with the **weak-* topology** induced by the family of seminorms $|\cdot|_\psi$ with $\psi \in \mathbb{H}$ defined by $|\phi|_\psi := |\phi(\psi)|$ for $\phi \in \mathbb{H}^\times$. Thus, $U \subseteq \mathbb{H}^\times$ is a neighborhood of $\phi \in \mathbb{H}^\times$ iff there are finitely many $\psi_k \in \mathbb{H}$ such that U contains all $\phi' \in \mathbb{H}^\times$ with $|\phi'(\psi_k) - \phi(\psi_k)| \le 1$ for all k. (The 1 can be replaced by any positive constant since the ψ_k can be arbitrarily scaled.) As a consequence, a net[1] of vectors $\phi_\ell \in \mathbb{H}^\times$ **converges** in the weak-* topology to the **weak-* limit** $\phi \in \mathbb{H}^\times$ iff $\phi_\ell(\psi) \to \phi(\psi)$ for all $\psi \in \mathbb{H}$. Because of (4.2), we may identify $\psi \in \mathbb{H}$ with the antilinear functional on $\mathbb{H}$, defined by

$$\psi(\phi) := \langle \phi, \psi \rangle \quad \text{for } \phi \in \mathbb{H}. \tag{4.3}$$

This definition turns $\mathbb{H}$ canonically into a subspace of $\mathbb{H}^\times$.

Example 4.1.1. The vector space $M(Z)$ of complex-valued functions $\psi : Z \to \mathbb{C}$ with finite support is with the inner product

$$\langle \phi, \psi \rangle := \sum_{z \in Z} \overline{\phi(z)} \psi(z),$$

a Euclidean space. The antidual $M(Z)^\times$ is the space of all complex-valued functions $\psi : Z \to \mathbb{C}$, with

$$\psi(\phi) := \sum_{z \in Z} \overline{\phi(z)} \psi(z).$$

Weak-* convergence in $M(Z)$ is just pointwise convergence.

---

1 All limits are formulated in terms of nets indexed by a directed set rather than sequences indexed by nonnegative integers, to cover the possibility of nonseparable spaces. In a separable Hilbert space, net convergence and sequence convergence are equivalent. In general, there is a difference, and nets are needed to obtain the correct topology.
For those not familiar with nets—they are generalizations of sequences defining the appropriate form of the limit in the nonseparable case. In the separable case, nets can always be replaced by sequences. Thus, readers will grasp the main content if, on first reading, they simply think of nets as being sequences.

Proposition 4.1.2.

(i) *Every $\psi \in \mathbb{H}^\times$ is the weak-* limit of a net of vectors from $\mathbb{H}$.*

(ii) *For every weak-* continuous antilinear functional Ψ on a subspace V of $\mathbb{H}^\times$, there is a $\psi \in \mathbb{H}$ such that*

$$\Psi(\phi) = \overline{\phi(\psi)} \quad \text{for } \phi \in V.$$

(iii) *Every weak-* continuous antilinear functional on $\mathbb{H}$ has a unique extension to a weak-* continuous antilinear functional on $\mathbb{H}^\times$.*

Proof. (i) For any finite-dimensional subspace V of $\mathbb{H}$, there is a unique $\psi_V \in V$ such that $\psi(\phi) = \langle \phi, \psi_V \rangle$ for all $\phi \in V$. The collection of finite-dimensional subspaces form a directed set under inclusion; hence the ψ_V form a net. The net converges to ψ in the weak-* topology since for all $\phi \in \mathbb{H}$,

$$(\psi - \psi_V)(\phi) = \langle \phi, \psi - \psi_V \rangle = \langle \phi, \psi \rangle - \langle \phi, \psi_V \rangle \to 0.$$

(ii) By continuity, there is a neighborhood N of zero such that

$$\big|\Psi(\phi)\big| \le 1 \quad \text{for all } \phi \in N.$$

By definition of the weak-* topology, there are $\psi_1, \dots, \psi_n \in \mathbb{H}$ such that N contains all $\phi \in \mathbb{H}$ with $|\phi(\psi_k)| \le 1$ for $k = 1, \dots, n$. Let $A : V \to \mathbb{C}^n$ be the linear mapping with $(A\phi)_k := \phi(\psi_k)$ for all k. If $A\phi = 0$ and $\varepsilon > 0$, then $\varepsilon^{-1}\phi \in N$; hence $|\Psi(\varepsilon^{-1}\phi)| \le 1$. Thus, $|\Psi(\phi)| \le \varepsilon$ for all $\varepsilon > 0$, giving $\Psi(\phi) = 0$.

This implies that $f(A\phi) := \Psi(\phi)$ defines an antilinear functional f on the range of A. We may extend f to an antilinear functional on $\mathbb{C}^n$. This has the form $f(x) = u^T \overline{x}$ with suitable $u \in \mathbb{C}^n$. Now $\psi := \sum u_k \psi_k$ is in $\mathbb{H}$ and satisfies

$$\phi(\psi) = \phi\Big(\sum u_k \psi_k\Big) = \sum \overline{u}_k \phi(\psi_k) = \sum \overline{u}_k (A\phi)_k,$$
$$\overline{\phi(\psi)} = \sum u_k \overline{(A\phi)_k} = u^T \overline{A\phi} = f(A\phi) = \Psi(\phi).$$

(iii) This follows from (ii) for $V = \mathbb{H}$. □

We define the **adjoint** ψ^* of $\psi \in \mathbb{H}$ to be the linear functional on $\mathbb{H}^\times$ that maps $\phi \in \mathbb{H}^\times$ to

$$\psi^* \phi := \phi(\psi),$$

and the **adjoint** ψ^* of $\psi \in \mathbb{H}^\times$ to be the linear functional on $\mathbb{H}$ that maps $\phi \in \mathbb{H}$ to

$$\psi^* \phi := \overline{\psi(\phi)}.$$

As a consequence,

$$\langle \phi, \psi \rangle = \phi^* \psi \quad \text{for } \phi, \psi \in \mathbb{H}.$$

Moreover, if $\phi, \psi \in \mathbb{H}^\times$ and one of them is in $\mathbb{H}$, then

$$\overline{\psi^* \phi} = \phi^* \psi.$$

Corollary 4.1.3.
(i) *Every linear mapping $f : \mathbb{H} \to \mathbb{C}$ can be written in the form $f = \phi^*$ for some $\phi \in \mathbb{H}^\times$.*
(ii) *Every weak-* continuous linear functional f on $\mathbb{H}^\times$ can be written in the form $f = \phi^*$ for some $\phi \in \mathbb{H}$.*

Proof. (i) The mapping $\phi : \mathbb{H} \to \mathbb{C}$ defined by $\phi(\psi) := \overline{f\psi}$ is antilinear; hence $\phi \in \mathbb{H}^\times$. Since $f\psi = \overline{\phi(\psi)} = \phi^* \psi$, we conclude that $f = \phi^*$.
(ii) Follows in the same way from Proposition 4.1.2(ii). □

We equip $\mathbb{H}$ with the **strict topology**, the locally convex topology, in which all antilinear (and hence all linear) functionals are continuous. Therefore, * is an anti-isomorphism from $\mathbb{H}^\times$ to the space of all linear functionals on $\mathbb{H}$, the **dual** of $\mathbb{H}$ with respect to the strict topology.

Example 4.1.4. In the Euclidean space $M(Z)$ defined in Example 4.1.1, a net ψ_ℓ in $M(Z)$ converges to $\psi \in M(Z)$ in the strict topology iff ψ is the weak-* limit, and there is a finite subset S of Z such that $\psi_\ell(z) = \psi(z)$ for all $z \in Z \setminus S$.

4.2 Norm and completion of a Euclidean space

Proposition 4.2.1. *The **Euclidean norm** $\|\psi\|$ defined on a Euclidean space $\mathbb{H}$ by*

$$\|\psi\| := \sqrt{\psi^* \psi}$$

*is positive when $\psi \neq 0$. It satisfies for $\phi, \psi \in \mathbb{H}$ the **Cauchy–Schwarz inequality***

$$|\phi^* \psi| \leq \|\phi\| \|\psi\|, \tag{4.4}$$

*and the **triangle inequality***

$$\|\phi + \psi\| \leq \|\phi\| + \|\psi\|, \tag{4.5}$$

and for $\lambda \in \mathbb{C}$ the relation

$$\|\lambda\psi\| = |\lambda| \|\psi\|.$$

Proof. (4.4) holds for $\psi = 0$. For $\psi \neq 0$ and $\beta := \psi^* \phi / \|\psi\|^2$,

$$0 \leq (\phi - \beta\psi)^* (\phi - \beta\psi) = \|\phi\|^2 - 2\,\mathrm{Re}(\beta \phi^* \psi) + |\beta|^2 \|\psi\|^2 = \|\phi\|^2 - |\phi^* \psi|^2 / \|\psi\|^2,$$

so that (4.4) holds also in this case. The triangle inequality now follows from

$$\|\phi + \psi\|^2 = \|\phi\|^2 + 2\,\mathrm{Re}\,\phi^* \psi + \|\psi\|^2 \leq \|\phi\|^2 + 2\|\phi\|\|\psi\| + \|\psi\|^2 = \big(\|\phi\| + \|\psi\|\big)^2.$$

The final equation is obvious. □

A mapping $f : \mathbb{H} \to \mathbb{C}$ is called **bounded** if there is a constant C such that $|f(\psi)| \leq C\|\psi\|$ for all $\psi \in \mathbb{H}$. A **Cauchy net** in $\mathbb{H}$ consists of a net ψ_ℓ in $\mathbb{H}$, such that for every $\varepsilon > 0$ there is an index N such that $\|\psi_j - \psi_k\| \leq \varepsilon$ for $j, k \geq N$. It is called **bounded** if $\sup_\ell \|\psi_\ell\| < \infty$. A **Hilbert space** is a Euclidean space containing with each bounded Cauchy net its weak-* limit.

Theorem 4.2.2. *The set $\overline{\mathbb{H}}$ of all bounded antilinear functionals on $\mathbb{H}$ is a Hilbert space, and we have*

$$\mathbb{H} \subseteq \overline{\mathbb{H}} \subseteq \mathbb{H}^\times. \tag{4.6}$$

If a net ψ_ℓ in $\overline{\mathbb{H}}$ has a weak- limit $\psi \in \overline{\mathbb{H}}$, then $\|\psi_\ell - \psi\| \to 0$.*

Proof. Clearly $\overline{\mathbb{H}}$ is a subspace of $\mathbb{H}^\times$. The Cauchy–Schwarz inequality says that, as an antilinear functional, $\psi \in \mathbb{H}$ is bounded. Thus, $\overline{\mathbb{H}}$ contains $\mathbb{H}$, and (4.6) holds.

Now suppose that $\phi, \psi \in \overline{\mathbb{H}}$ and that $\phi = \lim \phi_j$, $\psi = \lim \psi_j$ for nets with $\phi_j, \psi_j \in \mathbb{H}$. Then

$$|\phi_j^* \psi_j - \phi_k^* \psi_k| = |(\phi_j - \phi_k)^* \psi_j + \phi_k^*(\psi_j - \psi_k)| \leq \|\phi_j - \phi_k\|\|\psi_j\| + \|\phi_k\|\|\psi_j - \psi_k\|$$

converges to zero as $j, k \to \infty$. Therefore, the $\phi_j^* \psi_j$ form a Cauchy net and the limit

$$\phi^* \psi := \lim_\ell \langle \phi_\ell, \psi_\ell \rangle \tag{4.7}$$

exists. A similar argument shows that the limit is independent of the choice of the nets. We take (4.7) as the definition of the inner product in $\overline{\mathbb{H}}$. It is easy to see that the inner product is Hermitian and linear in the second argument. Therefore, $\overline{\mathbb{H}}$ is a Euclidean space. In particular, Proposition 4.2.1 applies with $\overline{\mathbb{H}}$ in place of $\mathbb{H}$.

To show completeness, let ϕ_ℓ be a bounded Cauchy net in $\overline{\mathbb{H}}$. Then for every $\psi \in \mathbb{H}$,

$$|\phi_\ell^* \psi - \phi_k^* \psi| = |(\phi_\ell - \phi_k)^* \psi| \leq \|\phi_\ell - \phi_k\|\|\psi\| \to 0 \quad \text{for } k, \ell \to \infty.$$

Hence, the $\phi_\ell^* \psi$ form a Cauchy net in $\mathbb{C}$ and converge. Thus,

$$f(\psi) := \lim_{\ell \to \infty} \phi_\ell^* \psi$$

defines a map $f : \mathbb{H} \to \mathbb{C}$. Since for $\mu, \mu' \in \mathbb{C}$ and $\psi, \psi' \in \mathbb{H}$,

$$f(\mu\psi + \mu'\psi') - \mu f(\psi) - \mu' f(\psi') = \lim_{\ell \to \infty} (\phi_\ell^*(\mu\psi + \mu'\psi') - \mu\phi_\ell^* \psi - \mu' \phi_\ell^* \psi') = 0,$$

f is linear, and by Corollary 4.1.3(i), $f = \phi^*$ for some $\phi \in \mathbb{H}^\times$. Clearly, ϕ is the weak-* limit of the ϕ_ℓ. Since the Cauchy net is bounded, ϕ is bounded, too; hence $\phi \in \overline{\mathbb{H}}$.

To prove the final statement, we note that

$$\begin{aligned}\|\psi_\ell - \psi\|^2 &= (\psi_\ell - \psi)^*(\psi_\ell - \psi_m) + (\psi_\ell - \psi)^*(\psi_m - \psi)\\ &\le \|\psi_\ell - \psi\|\|\psi_\ell - \psi_m\| + (\psi_\ell - \psi)^*(\psi_m - \psi).\end{aligned}$$

The first term goes to zero due to the Cauchy property, and the second term goes to zero due to weak-* convergence. □

Corollary 4.2.3 (Riesz representation theorem). *For every norm-continuous linear functional f on $\overline{\mathbb{H}}$, there is a vector $\psi \in \overline{\mathbb{H}}$ such that*

$$f(\phi) = \psi^*\phi \quad \text{for all } \phi \in \overline{\mathbb{H}}. \tag{4.8}$$

Proof. The mapping $\psi : \overline{\mathbb{H}} \to \mathbb{C}$ defined by $\psi(\phi) := \overline{f(\phi)}$ for $\phi \in \mathbb{H}$ is antilinear, hence belongs to $\overline{\mathbb{H}}$. □

We call $\overline{\mathbb{H}}$ the **completion** of $\mathbb{H}$. If $\mathbb{H}$ is finite-dimensional, then $\mathbb{H} = \overline{\mathbb{H}} = \mathbb{H}^\times$ by standard arguments, and all topologies considered are equivalent. If $\mathbb{H}$ is infinite-dimensional, then we usually[2] have $\mathbb{H} \neq \overline{\mathbb{H}} \neq \mathbb{H}^\times$. For example, the space $\mathbb{H} := C([-1,1])$ of continuous functions on $[-1,1]$ has as antilinear functionals not only all elements of the Hilbert space $\overline{\mathbb{H}} = L^2([-1,1])$ of square-integrable functions on $[-1,1]$, but also all function evaluation maps, corresponding to distributions. All these are elements of the antidual $\mathbb{H}^\times$. In infinite dimensions, the norm topology in $\overline{\mathbb{H}}$ is weaker than the strict topology in $\mathbb{H}$, but stronger than the weak-* topology in $\mathbb{H}^\times$.

By now, $\phi^*\psi$ is defined whenever $\phi, \psi \in \mathbb{H}^\times$, and either one of the two is in $\mathbb{H}$, or both are in $\overline{\mathbb{H}}$. Thus, we have a partial binary operation * on $\mathbb{H}^\times$, called the **partial inner product** (**PIP**). It satisfies

$$\overline{\phi^*\psi} = \psi^*\phi. \tag{4.9}$$

Unless $\mathbb{H}$ is finite-dimensional, the partial inner product is not everywhere defined. (For example, in the antidual of $C([-1,1])$, the inner product of two delta distributions at the same point is not defined.) Proposition 4.1.2(i) implies that $\mathbb{H}$ is dense in $\mathbb{H}^\times$. In particular, $\mathbb{H}^\times$ is a positive definite **PIP space** in the sense of Antoine & Trapani [11].

Example 4.2.4. The completion $\overline{M(Z)}$ of the Euclidean space $M(Z)$, discussed in Examples 4.1.1 and 4.1.4, is the Hilbert space of functions $\psi : Z \to \mathbb{C}$ with countable support and finite $\sum_{z\in Z} |\psi(z)|^2$. The inner product of $\phi, \psi \in \overline{M(Z)}$ is given by the absolutely convergent countable sum $\phi^*\psi := \sum_{z\in Z} \overline{\phi(z)}\psi(z)$.

2 This applies except when $\mathbb{H}$ is already an infinite-dimensional Hilbert space, in which case $\mathbb{H} = \overline{\mathbb{H}} \neq \mathbb{H}^\times$.

4.3 Linear mappings between Euclidean spaces

An **isometry** from a Euclidean space U to a Euclidean space V is a linear map $A : U \to V$, such that

$$\|A\psi\| = \|\psi\| \quad \text{for all } \psi \in U.$$

Isometries are injective since $A\psi = 0$ implies $\|\psi\| = 0$, and hence $\psi = 0$. An **isomorphism** from U to V is a surjective isometry $A : U \to V$. Its inverse is an isomorphism from V to U. If such an isomorphism exists, the Euclidean spaces U and V are called **isometric** or **isomorphic**.

If U and V are (complex) topological vector spaces, we write $\mathrm{Lin}(U, V)$ for the vector space of all continuous linear mappings from U to V, and $\mathrm{Lin}\, U$ for $\mathrm{Lin}(U, U)$. We identify V with the space $\mathrm{Lin}(\mathbb{C}, V)$ via

$$\psi\alpha := \alpha\psi \quad \text{for } \alpha \in \mathbb{C},\ \psi \in V.$$

Proposition 4.3.1. *Let U and V be Euclidean spaces.*

(i) *For any linear map $A : U \to V^\times$, the mapping $A^*\phi : U \to \mathbb{C}$, defined for $\phi \in V$ by*

$$(A^*\phi)(\psi) := (A\psi)^*\phi \quad \textit{for } \psi \in U,$$

is an antilinear functional and defines an operator $A^* : V \to U^\times$ with

$$(A\psi)^*\phi = \psi^*(A^*\phi),$$

*called the **adjoint** of A.*

(ii) *Any linear map $A : U \to V^\times$ is continuous, that is, $A \in \mathrm{Lin}(U, V^\times)$.*

(iii) *The mapping * that maps A to A^* is an antilinear mapping from $\mathrm{Lin}(U, V^\times)$ to $\mathrm{Lin}(V, U^\times)$ and satisfies*

$$A^{**} = A.$$

Proof. (i) is obvious.

(ii) We need to show that for every weak-$\star$ neighborhood N of 0 in $V^\times$, there is a strict neighborhood M of 0 in U such that $A\psi \in N$ for all $\psi \in M$. By definition of the weak-$\star$ topology, there are $\phi_1, \dots, \phi_n \in V$ such that N contains all $\phi \in V^\times$ with $|\phi(\phi_k)| \le 1$ for $k = 1, \dots, n$. The set M of all $\psi \in U$ with $|\psi(A^*\phi_k)| \le 1$ for $k = 1, \dots, n$ is a strict neighborhood of 0 in U and has the required property.

(iii) The dependence of $A^*\phi$ on ϕ is linear; thus the adjoint A^* is a linear operator. By (i), $A^* \in \mathrm{Lin}(V, U^\times)$. Corollary 4.1.3(ii) gives $V^{\times\times} = V$ and $U^{\times\times} = U$. Hence, $A^* : V \to U^\times$ is given by $A^*\psi(\phi) = (A\phi)^*\psi$ for all $\psi \in V$ and $\phi \in U$. Thus, we have for all $\phi \in U$ and $\psi \in V$,

$$A^{**}\phi(\psi) = \big(A^*\psi\big)^*\phi = A\phi(\psi),$$

which implies that $A^{**} = A$. □

Since $V \subseteq V^\times$, the adjoint is also defined for $A \in \mathrm{Lin}(U, V)$ and then makes sense as a mapping $A^* \in \mathrm{Lin}(V^\times, U^\times)$, and we have

$$A^*B^* = (BA)^* \quad \text{if } A \in \mathrm{Lin}(U, V), \quad B \in \mathrm{Lin}(V, W^\times).$$

We write

$$\mathrm{Lin}^\times \mathbb{H} := \mathrm{Lin}(\mathbb{H}, \mathbb{H}^\times)$$

for the vector space of continuous linear operators from a Euclidean space $\mathbb{H}$ to its antidual. Since $\mathbb{H}^{\times\times} = \mathbb{H}$ by Proposition 4.1.2(v), we conclude:

Corollary 4.3.2. *If $A \in \mathrm{Lin}^\times \mathbb{H}$, then $A^* \in \mathrm{Lin}^\times \mathbb{H}$, and we have*

$$\phi^* A\psi = (\phi^* A)\psi = \phi^*(A\psi) = (A^*\phi)^*\psi \quad \text{for } \phi, \psi \in \mathbb{H}. \tag{4.10}$$

Thus, $\phi^* A\psi$ defines a sesquilinear form on $\mathbb{H}$.

Here ϕ^* is treated as the adjoint $\phi^* : \mathbb{H}^\times \to \mathbb{C}$ of $\phi : \mathbb{C} \to \mathbb{H}$ under the identification $V = \mathrm{Lin}(\mathbb{C}, V)$. We call $A \in \mathrm{Lin}^\times \mathbb{H}$ **Hermitian** if $A^* = A$; then $\overline{\phi^* A\psi} = \psi^* A\phi$, so that the associated sesquilinear form is Hermitian.

4.4 Functions of positive type

A complex $n \times n$ matrix G is **Hermitian** if $\overline{G}_{jk} = G_{kj}$ for $j, k = 1, \dots, n$, **positive semidefinite** if $u^* Gu \geq 0$ for all $u \in \mathbb{C}^n$, and **conditionally semidefinite** if $u^* Gu \geq 0$ for all $u \in \mathbb{C}^n$ with $\sum_k u_k = 0$.

Let Z be a nonempty set. We call a function $F : Z \times Z \to \mathbb{C}$ **of positive type** (that is, **conditionally positive**) over Z if, for every finite sequence $z_1, \dots, z_n$ in Z, the **Gram matrix** of $z_1, \dots, z_n$, that is, the $n \times n$-matrix G with entries

$$G_{jk} = F(z_j, z_k), \tag{4.11}$$

is Hermitian and positive semidefinite (that is, conditionally semidefinite). In particular, every function of positive type is conditionally positive.

The basic intuition for the above definition comes from the following examples. (Note that z and z' are unrelated points.)

Proposition 4.4.1. *Let Z be a subset of a Euclidean space $\mathbb{H}$. Then the functions $F, F', F'' : Z \times Z \to \mathbb{C}$, defined by*

$$F(z, z') := z^* z', \quad F'(z, z') := z'^* z, \quad F''(z, z') := \mathrm{Re}\, z^* z',$$

are of positive type.

Proof. Let G, G', G'' be the Gram matrices computed with F, F', F'', respectively. Clearly, G is Hermitian; it is positive semidefinite since

$$u^* Gu = \sum_{j,k} \overline{u}_j z_j^* z_k u_k = \left\| \sum_k z_k u_k \right\|^2 \geq 0.$$

$G' = \overline{G}$ and $G'' = \frac{1}{2}(G + \overline{G})$ are easily seen to be Hermitian and positive semidefinite, too. □

The Moore–Aronszejn theorem (Theorem 4.6.1) provides a converse of Proposition 4.4.1.

Proposition 4.4.2. *If $F : Z \times Z \to \mathbb{C}$ is conditionally positive. Then, for any function $f : Z \to \mathbb{C}$ and any $\gamma \geq 0$, the function $\widetilde{F} : Z \times Z \to \mathbb{C}$, defined by*

$$\widetilde{F}(z, z') := \overline{f(z)} + f(z') + \gamma F(z, z') \quad \text{for } z, z' \in Z, \tag{4.12}$$

is conditionally positive.

Proof. Let $G, \widetilde{G}$ be the Gram matrices computed with F and $\widetilde{F}$, respectively. Clearly, $\widetilde{G}$ is Hermitian, and

$$\widetilde{G}_{jk} = \overline{f(z_j)} + f(z_k) + \gamma G_{jk};$$

hence $\sum_\ell u_\ell = 0$ implies

$$u^* \widetilde{G} u = \sum_{j,k} \overline{u}_j (\overline{f(z_j)} + f(z_k) + \gamma G_{jk}) u_k = \gamma \sum_{j,k} \overline{u}_j G_{jk} u_k = \gamma u^* G u \geq 0.$$

Thus, $\widetilde{G}$ is conditionally semidefinite. □

Proposition 4.4.3. *Let Z be a subset of a Euclidean space $\mathbb{H}$. Then, for any function $g : Z \to \mathbb{C}$, the function $\widetilde{F} : Z \times Z \to \mathbb{C}$, defined by*

$$\widetilde{F}(z, z') := \overline{g(z)} + g(z') - \|\overline{z} - z'\|^2 \quad \text{for } z, z' \in Z, \tag{4.13}$$

is conditionally positive.

Proof. This follows from Propositions 4.4.1 and 4.4.2 since $\widetilde{F}(z, z') = \overline{f(z)} + f(z') + F''(z, z')$, where $f(z) = g(z) - \|z\|^2$. □

For appropriate converses of Propositions 4.4.2 and 4.4.3, see the theorems by Schoenberg and by Menger in Section 4.9.

4.5 Constructing functions of positive type

In this section, we discuss a toolkit for the construction of such explicit functions of positive type from simpler ingredients. We provide a number of constructions that allow one to verify positivity properties. For further constructions and numerous examples in the form of exercises see Berg et al. [34].

Proposition 4.5.1. *For every family ϕ_z ($z \in Z$) of vectors ϕ_z in a Euclidean vector space $\mathbb{H}$, the function F, defined by*

$$F(z,z') := \langle \phi_z, \phi_{z'} \rangle,$$

is of positive type.

Proof. The corresponding matrix G from (4.11) is clearly Hermitian, and

$$x^* G x = \sum_{j,k} \overline{x}_j G_{jk} x_k = \sum_{j,k} \overline{x}_j \langle \phi_{z_j}, \phi_{z_k} \rangle x_k = \left| \sum_k x_k \phi_{z_k} \right|^2 \geq 0. \qquad \square$$

Basic examples of functions of positive type arise from the above constructions by choosing the family of ϕ_z in such a way that their inner products can be expressed in closed form. Others come from a number of constructions, which modify or combine functions of positive type.

Proposition 4.5.2.

(i) *Every positive semidefinite Hermitian form on a complex vector space Z is of positive type.*

(ii) *If F is of positive type over Z and $Y \subseteq Z$, then the **restriction** $F|_Y$ of F to $Y \times Y$ is of positive type.*

(iii) *If F_0 is of positive type over Z_0 and $u : Z \to Z_0$, then*

$$F(z,z') := F_0(u(z), u(z'))$$

is of positive type.

(iv) *If F is of positive type over Z, $\gamma > 0$, and $v : Z \to \mathbb{C}$, then*

$$F'(z,z') := \gamma \overline{v(z)} F(z,z') v(z')$$

*is of positive type. In particular, if $F(z,z) > 0$ for all z, then the **normalization** F_{norm} of F, defined by*

$$F_{\mathrm{norm}}(z,z') := \frac{F(z,z')}{\sqrt{F(z,z) F(z',z')}},$$

is of positive type, and satisfies $F_{\mathrm{norm}}(z,z) = 1$ for all z.

(v) *If L is a countable set and each F_ℓ ($\ell \in L$) is of positive type over Z, then, for arbitrary positive weights w_ℓ, for which*

$$F(z,z') := \sum_{\ell\in L} w_\ell F_\ell(z,z')$$

is everywhere defined, F is of positive type.

(vi) *Let Z be the disjoint union of a family of sets Z_ℓ, indexed by $\ell \in L$. If $F_\ell : Z_\ell \times Z_\ell \to \mathbb{C}$ is of positive type for all $\ell \in L$, then the function $F : Z \times Z \to \mathbb{C}$ (defined by*

$$F(z,z') := \begin{cases} F_\ell(z,z') & \text{if } z,z' \in Z_\ell, \\ 0 & \text{otherwise}) \end{cases}$$

is of positive type.

(vii) *If the F_ℓ ($\ell = 0,1,2\ldots$) are of positive type over Z and the limit*

$$F(z,z') := \lim_{\ell\to\infty} F_\ell(z,z')$$

exists for $z,z' \in Z$, then F is of positive type.

(viii) *If μ is a positive measure on a set L and each F_ℓ ($\ell \in L$) is of positive type over Z, then*

$$F(z,z') := \int_L d\mu(\ell) F_\ell(z,z'),$$

if everywhere defined, is of positive type.

(ix) *Let $F_0 : Z_0 \times Z_0 \to \mathbb{C}$ be of positive type, let $d\mu$ be a positive measure on a set L. If $u : Z \times L \to Z_0$ is such that the integral*

$$F(z,z')(\ell) := \int_L d\mu(\ell') F_0(u(z,\ell), u(z',\ell'))$$

exists for all $\ell \in L$ and $z,z' \in Z$, then F is of positive type on Z.

Proof. (i)–(viii) are straightforward, and (ix) follows from (iii) and (viii). □

Note that many examples of interest are analytic in the second argument. Unfortunately, this property does not persist under normalization as in Proposition 4.5.2(iv).

It is easily checked that all constructions of Proposition 4.5.2 produce conditionally positive functions when the ingredients are only required to be conditionally positive rather than of positive type.

Theorem 4.5.3 (Schur [271]).

(i) *If F_1 is of positive type on Z_1 and F_2 is of positive type on Z_2, then*

$$F((z_1,z_2),(z_1',z_2')) := F_1(z_1,z_1')F_2(z_2,z_2')$$

is of positive type on $Z = Z_1 \times Z_2$.

(ii) *If F_1 and F_2 are of positive type, then the pointwise product*

$$F(z,z') := F_1(z,z')F_2(z,z')$$

is of positive type.

Proof. (i) For $t = 1,2$, the Gram matrix G_t of $z_{t1},\dots,z_{tn}$, computed with respect to F_t is positive semidefinite, hence has a Cholesky factorization $G_t = R_t^* R_t$. The Gram matrix of $(z_{11},z_{21}),\dots,(z_{1n},z_{2n})$, computed with respect to F, has entries

$$\begin{aligned} G_{jk} = G_{1jk}G_{2jk} &= \Big(\sum_\ell \overline{R}_{1\ell j}R_{1\ell k}\Big)\Big(\sum_m \overline{R}_{2mj}R_{2mk}\Big) \\ &= \sum_{\ell,m} \overline{R}_{1\ell j}\overline{R}_{2mj}R_{1\ell k}R_{2mk}, \end{aligned}$$

so that

$$u^* Gu = \sum_{j,k} \overline{u}_j G_{jk} u_k = \sum_{\ell,m} \Big|\sum_j u_j R_{1\ell j} R_{2mj}\Big|^2 \geq 0.$$

Thus, G is positive definite; proving that F is of positive type.

(ii) follows from (i) and Proposition 4.5.2(iii) by mapping to the diagonal. □

Theorem 4.5.4.

(i) *All pointwise powers*

$$F^n(z,z') := F(z,z')^n \quad (n = 1,2,\dots)$$

of a function F of positive type are of positive type.

(ii) *If F is of positive type, then for any $\beta \geq 0$, the function F_β (defined by*

$$F_\beta(z,z') := e^{\beta F(z,z')})$$

is of positive type, too.

(iii) *Write $B(0;1) := \{x \in \mathbb{C} \mid |x| < 1\}$ for the open complex unit disk. If F is of positive type and $|F(z,z')| < c < \infty$ for all $z,z' \in Z$, then*

$$F_{\text{inv}}(z,z') := \frac{1}{c - F(z,z')}$$

is of positive type, too. (This is related to Nevanlinna–Pick interpolation theory; see Agler & McCarthy [1].*)*

Proof. (i) follows from Theorem 4.5.3(ii) by induction. (ii) and (iii) then follow from Proposition 4.5.2(v) since $e^x = \sum_0^\infty \frac{x^n}{n!}$ for $x \in \mathbb{C}$ and $\frac{1}{c-x} = \sum_0^\infty \frac{x^n}{c^{n+1}}$ for $|x| < c$, and constant functions with positive values are of positive type. □

This theorem is related to the Berezin–Wallach set discussed in Section 4.10.

4.6 The Moore–Aronszajn theorem

The remaining sections of this chapter provide, in a physics-oriented terminology and with full proofs, a self-contained synopsis (and sometimes slight generalization) of a number of classical results from the literature about functions of positive-type and reproducing kernel Hilbert spaces.

This section discusses the Moore–Aronszajn theorem, telling how to reconstruct a Hilbert space from a spanning set of vectors, whose inner product is known, and the properties that must be satisfied for arbitrarily assigned formal inner products to produce a Hilbert space. It leads in the next chapter to the existence of the quantum space of a coherent space, hence is of fundamental importance. Many other constructions in algebraic quantum mechanics (for example, the GNS construction of Hilbert spaces from states in C^*-algebras, or the Wightman reconstruction theorem in quantum field theory) are direct consequences of this theorem.

The Moore–Aronszajn theorem is due to Aronszajn [16] (1943), who attributed[3] it to Moore (1935).

Theorem 4.6.1 (Moore, Aronszajn). *Let $K : Z \times Z \to \mathbb{C}$ be of positive type. Then there is a unique Hilbert space $\overline{\mathbb{Q}}$ of complex-valued functions on Z with the Hermitian inner product $\langle \cdot, \cdot \rangle$ (antilinear in the first component) such that the following properties hold:*
(i) *$\overline{\mathbb{Q}}$ contains the functions $q_z : Z \to \mathbb{C}$, defined for $z \in Z$ by*

$$q_z(x) := K(x, z) = \overline{K(z, x)}. \tag{4.14}$$

(ii) *The space $\mathbb{Q}$ of finite linear combinations of the q_z is dense in $\overline{\mathbb{Q}}$.*
(iii) *The following relations hold:*

$$\langle q_z, q_x \rangle = K(z, x), \tag{4.15}$$

$$\psi(z) = \langle q_z, \psi \rangle \quad \text{for all } \psi \in \overline{\mathbb{Q}}. \tag{4.16}$$

(iv) *For each $z \in Z$, the linear functional ι_z, defined by*

$$\iota_z \psi := \psi(z), \tag{4.17}$$

is continuous.

3 Aronszajn [15, Théorème 2] states the theorem and gives a detailed proof (in French), but his later English paper [16] states the theorem on page 344 and attributes it to Moore. He cites Moore [186] (and a very short notice from 1916) on page 338, but the theorem does not seem to be in one of these references. (Moore discusses in Chapter III functions of positive type under the name *positive Hermitian matrices*—see the statement at the top of page 182—but does not construct a Hilbert space from them.) Faraut & Korányi [85, page 170] ascribes the theorem to Bergman [35] (1933), but the theorem does not seem to be there either. Kolmogorov [163, Lemma 2] (1941) contains the result for the special case, where Z is countable.

Proof. The vector space $\mathbb{Q}$ spanned by the q_z consists of all linear combinations

$$\widehat{f} := \sum_z f(z) q_z \tag{4.18}$$

with f in the space $\mathbb{F}$ of all maps $f : Z \to \mathbb{C}$ for which all but finitely many values $f(z)$ vanish. Thus, the sum is finite, and by (4.14), function values can be calculated by

$$\widehat{f}(x) = \sum_z f(z) q_z(x) = \sum_z K(x,z) f(z). \tag{4.19}$$

Since it might be possible that a function $\psi \in \mathbb{Q}$ can be written in several ways in the form (4.18), the definition of an inner product on $\mathbb{Q}$ requires some care. The mapping defined on $\mathbb{F} \times \mathbb{F}$ by

$$(g,f) := \sum_x \overline{g(x)} \widehat{f}(x) = \sum_{x,z} \overline{g(x)} K(x,z) f(z) \tag{4.20}$$

is a Hermitian form since

$$\overline{(g,f)} = \sum_{x,z} g(x) \overline{K(x,z) f(z)} = \sum_{x,z} \overline{f(z)} K(z,x) g(x) = (f,g).$$

Now

$$(g,f) = \overline{(f,g)} = \sum_x f(x) \overline{\widehat{g}(x)} = \sum_z \overline{\widehat{g}(z)} f(z). \tag{4.21}$$

If $\widehat{g} = \widehat{u}$ and $\widehat{f} = \widehat{v}$, then

$$\begin{aligned}(g,f) &= \sum_z \overline{\widehat{g}(z)} f(z) = \sum_z \overline{\widehat{u}(z)} f(z) = (u,f) \\ &= \sum_x \overline{u(x)} \widehat{f}(x) = \sum_x \overline{u(x)} \widehat{v}(x) = (u,v).\end{aligned}$$

Therefore, (g,f) depends only on the functions $\widehat{g}$ and $\widehat{f}$. Thus,

$$\langle \psi, \psi' \rangle := (g,f) \quad \text{if } \psi = \widehat{g},\ \psi' = \widehat{f}$$

defines a Hermitian form $\langle \cdot, \cdot \rangle$ on $\mathbb{Q}$ satisfying

$$\langle \widehat{g}, \widehat{f} \rangle = (g,f). \tag{4.22}$$

The function $g_z \in \mathbb{F}$ defined (for arbitrary but fixed $z \in Z$) by $g_z(x) = 1$ if $x = z$ and $g_z(x) = 0$ otherwise, satisfies

$$\widehat{g}_z = q_z \tag{4.23}$$

by (4.18), hence by (4.21),

$$\langle q_z, \widehat{f}\rangle = \langle \widehat{g}_z, \widehat{f}\rangle = (g,f) = \sum_x \overline{g_z(x)}\widehat{f}(x) = \widehat{f}(z).$$

Since by definition of $\mathbb{Q}$ any $\psi \in \mathbb{Q}$ can be written as $\psi = \widehat{f}$, we conclude (4.16). Specialization to $\psi = q_x$ and using (4.14) (with z and x interchanged) yield (4.15).

Since K is of positive type, $(f,f) \geq 0$ for all $f \in \mathbb{F}$. Therefore, the form is positive semidefinite on $\mathbb{F}$. In particular, the Cauchy–Schwarz inequality $|(f,f')|^2 \leq (f,f)(f',f')$ holds. It implies that $(f,f) = 0$ only if $(f,f') = 0 = (f',f)$ for all f', and (4.18) then shows that $\widehat{f}(z) = 0$ for all z. Hence, $\widehat{f} = 0$. Therefore, the Hermitian form $\langle\cdot,\cdot\rangle$ is positive definite, hence defines a Hermitian inner product on $\mathbb{Q}$. Thus, $\mathbb{Q}$ is a Euclidean space. The completion with respect to the norm

$$\|\psi\| := \sqrt{\langle\psi,\psi\rangle}$$

(which can be done constructively using Theorem 4.2.2) gives the desired Hilbert space, and a limiting argument shows that (4.16) holds in general: If $\psi \in \overline{\mathbb{Q}}$, there is a net of $\psi_j \in \mathbb{Q}$, converging to ψ in the norm, and

$$\big|\langle q_z,\psi\rangle - \psi_j(z)\big| = \big|\langle q_z,\psi\rangle - \langle q_z,\psi_j\rangle\big| = \big|\langle q_z,\psi-\psi_j\rangle\big| \leq \|q_z\|\|\psi-\psi_j\| \to 0.$$

Hence, $\psi(z) = \lim_j \psi_j(z) \to \langle q_z,\psi\rangle$.

(iv) Since $\iota_z\psi = \psi(z) = \langle q_z,\psi\rangle$, we have $\|\iota_z\| = \|q_z\|$. Thus, ι_z is bounded and hence continuous.

The uniqueness of $\overline{\mathbb{Q}}$ is clear from the construction. □

4.7 Reproducing kernel Hilbert spaces and Mercer's theorem

Proposition 4.7.1. *Let ψ_α ($\alpha \in I$) be an orthonormal basis for $\overline{\mathbb{Q}}$. Then*

$$K(z,w) = \sum_{\alpha\in I} \psi_\alpha(z)\overline{\psi_\alpha(w)}. \tag{4.24}$$

Proof. By the polarized version of the Parseval identity, Theorem 5.27 of Folland [90], we have

$$q_w = \sum_{\alpha\in I}\langle\psi_\alpha, q_w\rangle\psi_\alpha = \sum_{\alpha\in I}\overline{\psi_\alpha(w)}\psi_\alpha$$

for all $w \in Z$. Hence, for all $z,w \in Z$,

$$K(z,w) = \langle q_z,q_w\rangle = \left\langle q_z, \sum_{\alpha\in I}\overline{\psi_\alpha(w)}\psi_\alpha\right\rangle = \sum_{\alpha\in I}\langle q_z,\psi_\alpha\rangle\overline{\psi_\alpha(w)} = \sum_{\alpha\in I}\psi_\alpha(z)\overline{\psi_\alpha(w)},$$

which implies (4.24). □

A **reproducing kernel Hilbert space** is a Hilbert space $\mathbb{K}$ of functions on a set Z together with a **reproducing kernel** $K : Z \times Z \to \mathbb{C}$ such that the functions k_z ($z \in Z$), defined by

$$k_z(x) := K(x, z), \tag{4.25}$$

span a space dense in $\mathbb{K}$ and satisfy

$$\psi(z) = k_z^* \psi \quad \text{for all } \psi \in \mathbb{K},\ z \in Z. \tag{4.26}$$

Thus (4.16) says that $\overline{\mathbb{Q}}$ is a reproducing kernel Hilbert space with reproducing kernel K and $k_z = q_z$.

Proposition 4.7.1 is related to Mercer's theorem (Mercer [185]), which represents certain reproducing kernels by an infinite sum of the form

$$K(z, w) = \sum_{\alpha \in I} \lambda_\alpha \phi_\alpha(z) \overline{\phi_\alpha(w)},$$

with positive real numbers λ_α and functions ϕ_α satisfying additional properties. Precise statements of Mercer's theorem and its generalizations (for example, Ferreira & Menegatto [86]) require additional structure on Z and K concerning measurability and continuity, hence are not valid in the generality discussed here. A discussion of measure theoretic properties of coherent states and associated overcompleteness relations will be given elsewhere.

4.8 Theorems by Bochner and Kreĭn

Bochner [38, Satz 4] proved the following optimality result for q_x:

Theorem 4.8.1 (Bochner). *Let $K : Z \times Z \to \mathbb{C}$ be of positive type, and let $\mathbb{Q}$ be the space constructed in the Moore–Aronszejn theorem (Theorem 4.6.1). If $x \in Z$ satisfies $K(x, x) \neq 0$, then*

$$\min\{\psi^* \psi \mid \psi \in \mathbb{Q},\ \psi(x) = \alpha\} = \frac{|\alpha|^2}{K(x, x)}.$$

The minimum is attained just for $\psi = \frac{\alpha}{K(x,x)} q_x$. In particular, if $\alpha = K(x, x)$, then the minimum is attained precisely at q_x.

Proof. This is trivial for $\alpha = 0$. For $\alpha \neq 0$, we may rescale the assertion. Thus, it is enough to prove the case $\alpha = K(x, x)$. In this case,

$$\begin{aligned} \psi^* \psi &= \langle \psi - q_x, \psi - q_x \rangle + 2 \operatorname{Re} \langle q_x, \psi \rangle - \langle q_x, q_x \rangle \\ &= \langle \psi - q_x, \psi - q_x \rangle + 2 \operatorname{Re} \psi(x) - K(x, x) \\ &= \|\psi - q_x\|^2 + K(x, x) \geq K(x, x) = \alpha, \end{aligned}$$

with equality iff $\psi - q_x = 0$. □

Our next result, a variant of Kreĭn [167], characterizes which functions $\psi \in \mathbb{Q}^\times$ belong already to the Hilbert space $\overline{\mathbb{Q}}$.

Theorem 4.8.2 (Kreĭn). *Let $K : Z \times Z \to \mathbb{C}$ be of positive type and $\psi : Z \to \mathbb{C}$. Define the function $K_\varepsilon : Z \times Z \to \mathbb{C}$ by*

$$K_\varepsilon(z, z') := K(z, z') - \varepsilon\psi(z)\overline{\psi(z')}.$$

(i) *If $\psi \in \overline{\mathbb{Q}}$ and $0 < \varepsilon \le \|\psi\|^{-2}$, then K_ε is of positive type.*
(ii) *If K_ε is of positive type for some $\varepsilon > 0$, then $\psi \in \overline{\mathbb{Q}}$.*

Proof. (i) Hermiticity is obvious. To show that K_ε is of positive type, we need to show for any finite sequence of complex numbers u_k and points $z_k \in Z$ the nonnegativity of the sum

$$\sigma := \sum_{j,k} \overline{u}_j K_\varepsilon(z_j, z_k) u_k = \sum_{j,k} \overline{u}_j \langle q_{z_j}, q_{z_k}\rangle u_k - \varepsilon \sum_{j,k} \overline{u}_j \psi(z_j)\overline{\psi(z_k)} u_k,$$

where we used (4.15). Writing

$$q := \sum_k q_{z_k} u_k,$$

we find that

$$\langle \psi, q\rangle = \sum_k \langle \psi, q_{z_k}\rangle u_k = \sum_k \overline{\langle q_{z_k}, \psi\rangle} u_k = \sum_k \overline{\psi(z_k)} u_k;$$

hence

$$\sigma = \|q\|^2 - \varepsilon\big|\langle \psi, q\rangle\big|^2 \ge \|q\|^2 - \varepsilon\|\psi\|^2\|q\|^2 \ge 0.$$

(ii) In this case, with $\mathbb{F}$ and $\widehat{f}$ as in the proof of the Moore–Aronszejn theorem (Theorem 4.6.1), we consider the antilinear mapping $\Psi : \mathbb{F} \to \mathbb{C}$, defined by

$$\Psi(f) := \sum_{z \in Z} \overline{f(z)}\psi(z).$$

Since K_ε is of positive type, we have

$$\begin{aligned} 0 &\le \sum_{z,z'} \overline{f(z)} K_\varepsilon(z, z') f(z') = \sum_{z,z'} \overline{f(z)} K(z, z') f(z') - \varepsilon \sum_{z,z'} \overline{f(z)}\psi(z)\overline{\psi(z')} f(z') \\ &= (f, f) - \varepsilon\big|\Psi(f)\big|^2 = \|\widehat{f}\|^2 - \varepsilon\big|\Psi(f)\big|^2 \end{aligned}$$

by definition of K_ε, (4.20), (4.22), and the definition of Ψ. Therefore,

$$\big|\Psi(f)\big| \le \varepsilon^{-1/2}\|\widehat{f}\|.$$

In particular, $\widehat{f} = 0$ implies $\Psi(f) = 0$. Therefore, Ψ defines a unique antilinear mapping $\psi' : \mathbb{Q} \to \mathbb{C}$ with $\psi'(\widehat{f}) = \Psi(f)$ for all $f \in \mathbb{F}$. By the above, $|(\psi'\widehat{f})| \le \varepsilon^{-1/2}\|\widehat{f}\|$. Thus, ψ' is bounded. By Theorem 4.2.2, $\widehat{\Psi}$ belongs to

$$\widehat{\Psi}(\phi) = \langle \phi, \psi' \rangle \quad \text{for all } \psi \in \overline{\mathbb{Q}}.$$

Since by (4.23) and (4.16),

$$\psi(z) = \Psi(g_z) = \widehat{\Psi}(\widehat{g}_z) = \widehat{\Psi}(q_z) = \langle q_z, \psi' \rangle = \psi'(z) \tag{4.27}$$

for all $z \in Z$, we conclude that $\psi = \psi' \in \overline{\mathbb{Q}}$. □

4.9 Theorems by Schoenberg and Menger

In this section, we prove the promised converse of Propositions 4.4.2–4.4.3.

Theorem 4.9.1 (Schoenberg [267, p. 49]). *If F is conditionally positive, then the function P_a (defined for any $a \in Z$ by*

$$P_a(z, z') := F(z, z') - F(z, a) - F(a, z') + F(a, a)) \tag{4.28}$$

is of positive type. Conversely, if a map $F : Z \times Z \to \mathbb{C}$ is such that if P_a is of positive type for some *$a \in Z$, then F is conditionally positive.*

Proof. Let $G, \widetilde{G}$ be the Gram matrices of $z_1, \dots, z_n$ computed with F and P_a, respectively. Then

$$\widetilde{G} = G - g\mathbf{1}^* - \mathbf{1}g^* + \gamma\mathbf{1}\mathbf{1}^*,$$

where $\mathbf{1}$ is the all-one column vector, g the column vector with components $g_j := F(z_j, a)$, and $\gamma := F(a, a)$. The Gram matrix of $z_1, \dots, z_n, a$ computed with F is therefore

$$G' := \begin{pmatrix} G & g \\ g^* & \gamma \end{pmatrix}.$$

Now $v \in \mathbb{C}^{n+1}$ satisfies $\sum_j v_j = 0$ iff, for some $u \in \mathbb{C}^n$,

$$v = \begin{pmatrix} u \\ -s \end{pmatrix}, \quad s = \mathbf{1}^* u,$$

and then

$$\begin{aligned} v^* G' v &= \begin{pmatrix} u \\ -s \end{pmatrix}^* \begin{pmatrix} G & g \\ g^* & \gamma \end{pmatrix} \begin{pmatrix} u \\ -s \end{pmatrix} = u^* G u - u^* g s - \bar{s} g^* u + \gamma \bar{s} s \\ &= u^* (G - g\mathbf{1}^* - \mathbf{1}g^* + \gamma \mathbf{1}\,\mathbf{1}^*) u = u^* \widetilde{G} u. \end{aligned}$$

This shows that F is conditionally positive iff all $\widetilde{G}$ are positive semidefinite. That is, iff P_a is of positive type for some a and hence for all a. □

Theorem 4.9.2. *A map $F : Z \times Z \to \mathbb{C}$ is conditionally positive iff there is an embedding $z \to q_z$ of Z into a Euclidean space $\mathbb{H}$ such that*

$$F(z,z') = \overline{f(z)} + f(z') + q_z^* q_{z'} \tag{4.29}$$

holds for some $f : Z \to \mathbb{C}$.

Proof. (i) Suppose that F is conditionally positive. Fix a and define P_a by (4.28). By Theorem 4.9.1, P_a is of positive type. Hence, the Moore–Aronszejn theorem (Theorem 4.6.1) gives an embedding $z \to q_z$ into a Hilbert space such that

$$P_a(z,z') = q_z^* q_{z'}, \tag{4.30}$$

applying (4.15) of the theorem to P_a in place of K. The definition of P_a then implies

$$F(z,z') - F(z,a) - F(a,z') + F(a,a) = q_z^* q_{z'}.$$

Putting $z = z' = a$ gives $q_a^* q_a = 0$. Hence, $q_a = 0$. One now easily verifies that

$$D(z,z') := F(z,z') - q_z^* q_{z'}$$

satisfies $\overline{D(z,z')} = D(z',z)$ and

$$D(z,z') - D(z,a) - D(a,z') + D(a,a) = 0.$$

This implies that $D(z,z') = \overline{f(z)} + f(z')$ with

$$f(z) := D(a,z) - \frac{1}{2} D(a,a).$$

Therefore (4.29) holds.

(ii) Conversely, if (4.29) holds, then (4.30) and (4.28) imply that $P_a(z,z') = \langle q_z - q_a, q_{z'} - q_a \rangle$. Hence, P_a is of positive type. By Schoenberg's Theorem 4.9.1, F is conditionally positive. □

The following converse of Proposition 4.4.3 is related to results by Menger [184] in the context of characterizing metric spaces embeddable into a finite-dimensional real vector space.

Corollary 4.9.3. *A map $F : Z \times Z \to \mathbb{C}$, satisfying $F(z,z') = F(z',z)$ for $z, z' \in Z$, is conditionally positive iff there is an embedding $z \to q_z$ of Z into a real Euclidean space such that*

$$F(z,z') = g(z) + g(z') - \|q_z - q_{z'}\|^2 \tag{4.31}$$

holds for some $g : Z \to \mathbb{R}$.

Proof. If there is such an embedding, then F is conditionally positive by Proposition 4.4.3. Conversely, suppose that F is conditionally positive. Then $\frac{1}{2}F$ is also conditionally positive. By Theorem 4.9.2, there is an embedding $z \to q_z$ of Z into a complex Euclidean space $\mathbb{H}$ such that

$$\frac{1}{2}F(z,z') = \overline{f(z)} + f(z') + q_z^* q_{z'}$$

holds for some $f : Z \to \mathbb{C}$. Since we assumed $F(z,z') = F(z',z)$ for $z, z' \in Z$, $f(z)$ is real. Thus, the inner products $K(z,z') = q_z^* q_{z'}$ are also real, and the q_z span a real Euclidean space. Substitution of $g(z) = 2f(z) + \|q_z\|^2$ now shows that (4.31) holds. □

4.10 The Berezin–Wallach set

Many coherent products of interest have the exponential form discussed in the theorem that follows. In the case where F takes only finite values and is zero on the diagonal it is due to Schoenberg [266], to Herz [127, Proposition 6] where F takes only finite values, and to Horn [139] in the general case. The present proof is much shorter than Horn's.

To be able to formulate the results, we put

$$e^{-\infty} := 0,$$

and call a function $F : Z \times Z \to \mathbb{C} \cup \{-\infty\}$ **conditionally positive** if either (i) there is an equivalence relation $\equiv$ on Z such that F is conditionally positive on each equivalence class, and $F(z,z') = -\infty$ whenever $z \not\equiv z'$, or (ii) F takes only infinite values. This reduces to the original definition if the value $-\infty$ is not attained, which holds iff there is only one equivalence class.

Theorem 4.10.1.

(i) *If $F : Z \times Z \to \mathbb{C} \cup \{-\infty\}$ is conditionally positive, then (for all $\beta > 0$)*

$$K(z,z') := e^{\beta F(z,z')} \tag{4.32}$$

is of positive type.

(ii) *Let $F : Z \times Z \to \mathbb{C} \cup \{-\infty\}$. If there is a sequence of positive numbers β_k converging to 0 such that*

$$K_k(z,z') := e^{\beta_k F(z,z')}$$

is of positive type for all k, then F is conditionally positive.

Proof. (i) If F takes only finite values, then Theorem 4.9.1 shows that (for any $z_0 \in Z$), the function $\tilde{F}$, defined by

$$\tilde{F}(z,z') := \beta(F(z,z') - F(z,z_0) - F(z_0,z') + F(z_0,z_0)),$$

is of positive type. Theorem 4.5.4 therefore implies that $K(z,z') := e^{\bar{F}(z,z')}$ defines a function K of positive type. Rescaling this by Proposition 4.5.2(iv), we see that (4.32) is of positive type, too. If F takes infinite values only, K is identically zero, and hence of positive type. If F takes finite and infinite values, the previous argument may be applied to the restriction of K to each equivalence class, and shows that this restriction is of positive type. Then Proposition 4.5.2(vi) implies that K itself is of positive type.

(ii) We may assume, without loss of generality, that Z cannot be decomposed as in Proposition 4.5.2(vi).

Case 1: $K(z,z') \neq 0$ for all $z,z' \in Z$. We fix $a \in Z$ and use Proposition 4.5.2(iv) to rescale $K := e^{\beta F}$ (consistently for all β) such that all $K(a,z) = 1$. Hence, all $F(a,z)$ vanish. Theorem 4.9.1, applied with K_k in place of F, implies that the map $P_a : Z \times Z \to \mathbb{C}$, defined by

$$P_a(z,z') = K_k(z,z') - K_k(z,a) - K_k(a,z') + K_k(a,a) = K_k(z,z') - 1,$$

is of positive type. Therefore, the functions F_k, defined by

$$F_k(z,z') := \frac{K_k(z,z') - 1}{\beta_k} = \frac{e^{\beta_k F(z,z')} - 1}{\beta_k} = F(z,z') + \beta_k F(z,z')^2 + O(\beta_k)^2,$$

are also of positive type. Since $\beta_k \to 0$, $F_k(z,z') \to F(z,z')$ for $k \to \infty$. By Proposition 4.5.2(vii), F is of positive type. Undoing the scaling and using Theorem 4.9.1 now prove that F is conditionally positive.

Case 2: $K(z,z) = 0$ for some $z \in Z$. Then the positivity of the Gram matrix (4.11) for $n = 2$ implies that $K(z,z') = K(z',z) = 0$ for all $z' \in Z$, and the indecomposability assumed at the beginning of (ii) implies that $Z = \{z\}$ and K is identically zero. Thus, F takes only infinite values and is, therefore, conditionally positive.

Case 3: $K(z,z) \neq 0$ for all $z \in Z$, but $K(x,y) = 0$ for some $x,y \in Z$. By Proposition 4.5.2(iv), we may normalize K such that $K(z,z) = 1$ for all $z \in Z$. The Gram matrix

$$G = \begin{pmatrix} K(x,x) & K(x,y) & K(x,z) \\ K(y,x) & K(y,y) & K(y,z) \\ K(z,x) & K(z,y) & K(z,z) \end{pmatrix} = \begin{pmatrix} 1 & K(x,y) & K(x,z) \\ K(y,x) & 1 & K(y,z) \\ K(z,x) & K(z,y) & 1 \end{pmatrix}$$

of $x,y,z \in Z$ is Hermitian and positive semidefinite. Hence, its determinant is nonnegative:

$$0 \leq 1 - |K_k(x,z)|^2 - |K_k(y,z)|^2 = 1 - |K(x,z)|^{2\beta_k} - |K(y,z)|^{2\beta_k}.$$

Unless at least one of $K(x,z)$ or $K(y,z)$ vanishes, the two negative terms tend for $k \to \infty$ both to 1. Hence, the right hand side converges to –1. This holds for any z, whence Z can be split into two subsets X and Y such that $K(x,z) = 0$ for $z \in Y$, and $K(y,z) = 0$ for $z \in X$. By Hermiticity, $K(z,x) = 0$ for $z \in Y$, and $K(z,y) = 0$ for $z \in X$. Repeating the argument for all zeros constructible this way shows that Z decomposes as in Proposition 4.5.2(iii), a contradiction. □

If we want to discuss a possible generalization of Theorem 4.5.4 to other exponents, we need to assume that the power exists, which suggests to assume for $K(z, z')$ an exponential form.

The **Berezin–Wallach set** of a mapping $F : Z \times Z \to \mathbb{C} \cup \{-\infty\}$ is the set $W(F)$ of nonnegative real numbers β for which

$$K(z, z') := e^{\beta F(z,z')} \tag{4.33}$$

is of positive type. The **Berezin–Wallach set** of a coherent space is the set $W(F)$ where

$$F(z, z') := \log K(z, z'),$$

using the principal value of the logarithm and $\log 0 = -\infty$. (Thus, the Berezin–Wallach set of a coherent space always contains 1.)

This set was introduced by Wallach [293] in the context of representations of Lie groups. But already earlier, Berezin [33] computed the Berezin–Wallach set for the case when F is the Kähler potential of a Siegel domain. Indeed, in many cases of interest, Z is a so-called Kähler manifold, and F the associated Kähler potential; see, for example, Zhang et al. [316, Section VI]. For the Berezin–Wallach sets corresponding to Hermitian symmetric spaces see, for example, Faraut & Korányi [85, Section XIII.2].

Theorem 4.10.2.

(i) *The Berezin–Wallach set $W(F)$ is a closed set containing* 0.
(ii) *$W(F)$ contains with β and β' their sum, and hence all linear combinations with nonnegative integral coefficients.*
(iii) *If $W(F)$ contains an open set it contains all sufficiently large positive real numbers.*
(iv) *If F is conditionally positive, then $W(F)$ contains all nonnegative real numbers.*
(v) *If* 0 *is a limit point of $W(F)$, then F is conditionally positive.*

Proof. (i)–(iv) follow easily from Proposition 4.5.3(ii). (v) follows from Theorem 4.10.1. □

In the most interesting cases, the Berezin–Wallach set is of the form $\alpha\mathbb{M}_0 \cup [\beta, \infty]$ or $\alpha\mathbb{M}_0$, where $\alpha, \beta > 0$ and $\mathbb{M}_0$ denotes the set of nonnegative integers. In general, the Berezin–Wallach set may have a very complicated structure, already for Z with three elements only. FitzGerald & Horn [88] show that the Berezin–Wallach set of every finite coherent space Z with real, nonnegative coherent product contains the interval $[|Z| - 2, \infty[$.

5 Coherent spaces

The notion of a coherent space is a nonlinear version of the notion of a complex Euclidean space: The vector space axioms are dropped while the notion of inner product—now called a coherent product—is kept. Thus, length and angular properties of vectors in a Euclidean space (embodied in the inner product) are generalized in a similar way as, in the past, metric properties of Euclidean spaces were generalized to metric spaces, differential properties of Euclidean spaces were generalized to manifolds, and topological properties of Euclidean spaces were generalized to topological spaces.

This chapter defines concepts and basic theorems about coherent spaces. We treat the most basic aspects of coherent spaces, associated vector spaces, and their topology, starting from first principles. All proofs are carried out in detail.

To illustrate some of the connections to physics and complex analysis, we give a long but still very incomplete list of examples of coherent spaces. Some of these examples are very elementary and can be understood informally before reading the systematic exposition of the theory. Many more coherent spaces can be constructed by modifying given ones using the recipes from Section 4.5.

5.1 Motivation for coherent spaces

> *The scalar product may, in fact, be calculated more simply than by using wave functions.*
>
> Roy Glauber, 1963 [98, p. 2771]

In informal, traditional terms, a coherent space is roughly a set Z, whose elements label certain vectors, called **coherent states** of a Hilbert space. Its quantum space of Z is the subspace formed by the linear combinations of coherent states. However, one can characterize this situation independent of a Hilbert space setting. Many Euclidean spaces are described most simply and naturally in terms of a nice, small subset of coherent states, and all their properties can be investigated in terms of the associated coherent space. Thus, coherent spaces may also be viewed as a new, geometric way of working with concrete Hilbert spaces in which they are embeddable. In this context, coherent spaces give a natural geometric setting to the concept of coherent states.

In particular, while the study of most problems in traditional function spaces for applications rely heavily on measures and integration, the quantum spaces of coherent spaces—with an easily computable coherent product—can be studied efficiently without measures or integrations, in terms of the explicit coherent product and differentiation only; see Section 5.5 (below). This property, already noticed by Glauber [98], who coined the notion of a coherent state, makes many calculations easy that are difficult in Hilbert spaces, whose inner product is defined through a measure. As a consequence, coherent spaces combine the rich, often highly characteristic variety

https://doi.org/10.1515/9783110667387-005

of symmetries of traditional geometric structures with the computational tractability of traditional tools from numerical analysis and statistics.

A strength of the coherent space approach is that it makes many different things look alike, and stays close to actual computations. There are so many applications in physics and elsewhere, that pointing them all out in detail would require writing a whole book.

Coherent states and squeezed states in quantum optics, mean field calculations in statistical mechanics, Hartree–Fock calculations for the electronic states of atoms, semiclassical limits, integrable systems all belong here. Many computational techniques in quantum physics can be profitably phrased in terms of coherent spaces; see Section 6.8 (below).

As we shall see, coherent spaces abstract the essential geometric properties needed to define a reproducing kernel Hilbert space. Examples of reproducing kernels (that is, what in the present context are coherent products) were first discussed by Zaremba [312] in the context of boundary value problems and by Mercer [185] in the context of integral equations. The theory was systematically developed by Aronszajn [15, 16], Krein [167, 168], and others. For a history see Berg et al. [34] and Stewart [277].

Coherent spaces and reproducing kernel Hilbert spaces are mathematically almost equivalent concepts, and there is a vast literature related to the latter. Most relevant for the applications to quantum physics are the books by Perelomov [232], Neeb [188], and Neretin [189]. However, the emphasis in these books is quite different from the present exposition, as they are primarily interested in properties of the associated function spaces and group representations, whereas we are primarily interested in the geometry and symmetry properties and in computational tractability.

Coherent spaces have also close relations to Fock spaces, unitary group representations, and to many other fields of mathematics, statistics, physics, and engineering; see Section 5.10 (below).

5.2 Coherent spaces

Let Z be a nonempty set. A **coherent product** on Z is a function $K : Z \times Z \to \mathbb{C}$ of positive type.[1] A **coherent space**[2] is a nonempty set Z with a distinguished coherent

1 One obtains the more general concepts of **semicoherent products** and **semicoherent spaces** by weakening the requirement of having positive type to the requirement that the supremum $\mathrm{ns}(Z)$ of the number of negative eigenvalues of Gram matrices constructed from K is finite. Much of the subsequent theory remains valid, but the inner products need no longer be positive semidefinite and the quantum spaces discussed below become Pontryagin spaces with $\mathrm{ns}(Z)$ negative squares; see Alpay et al. [9]. In this book, this generalization is not considered further.

2 Note that some papers by Vourdas [290, 291] use the term "coherent space" for a different concept, also related to coherent states. Completely unrelated is the notion of coherent spaces used in logic.

product $K : Z \times Z \to \mathbb{C}$. We regard the same set with different coherent products as different coherent spaces.

Examples 5.2.1.

(i) Any subset Z of a Euclidean space is a coherent space with coherent product

$$K(z, z') := z^* z'.$$

(ii) Any subset Z' of a coherent space Z is again a coherent space, with the coherent product inherited from Z by restriction.

(iii) For practical applications, it is often important that the coherent products are given as explicit expressions $K(z, z')$, with which one can work analytically, or at least expressions, which can be efficiently approximated numerically. The easiest way to construct such expressions is by using one of the many constructions from Section 4.5.

In particular, coherent spaces generalize Euclidean spaces, and the coherent product $K(z, z')$ generalizes the Hermitian inner product $z^* z'$, but in general no linear structure is assumed on Z. This is similar to the way metric spaces generalize the distance in Euclidean spaces without keeping their linear structure.

Many interesting examples exists; some of these are discussed in Neumaier & Ghaani Farashahi [212]. We give one particularly important example.

Example 5.2.2. Let V be a Euclidean space. In a notation, where pairs are denoted by square brackets, we write

$$z := [z_0, \mathbf{z}] \in \mathbb{C} \times V$$

for the elements of $Z = \mathbb{C} \times V$ since

$$F(z, z') := \bar{z}_0 + z'_0 + \mathbf{z}^* \mathbf{z}' \tag{5.1}$$

is conditionally positive by Proposition 4.4.2. Theorem 4.10.1 implies that

$$K(z, z') := e^{\bar{z}_0 + z'_0 + \mathbf{z}^* \mathbf{z}'} \tag{5.2}$$

is a coherent product, with respect to which Z is a coherent space. We call this coherent space the **Klauder space** over V and denote it by $Kl[V]$. (For $V = \mathbb{C}$, the associated coherent states, the nonzero multiples of those discovered by Schrödinger [269], were first discussed in Klauder [157, p. 1062].)

Klauder spaces are discussed in more detail in Neumaier & Ghaani Farashahi [212]. In particular, this paper derives a coherent construction of creation annihilation operators together with their properties. As shown there, the quantum spaces of

Klauder spaces are essentially the **Fock spaces** introduced by Fock [89] in the context of **quantum field theory**. They were first presented by Segal [273] in a form equivalent to the above. The quantum space of $KL[\mathbb{C}^n]$ was systematically studied by Bargmann [24]. These coherent spaces are closely related to quantum field theory (Baez et al. [18], Glimm & Jaffe [99]) and the theory of Hida distributions in the white noise calculus for classical stochastic processes (Hida & Si [128], Hida & Streit [129], Obata [222]).

5.3 Quantum spaces

Let Z be a coherent space. A **quantum space** of Z is a Euclidean space $\mathbb{Q}(Z)$ spanned by (that is, consisting of all finite linear combinations of) a distinguished set of vectors $|z\rangle$ ($z \in Z$) satisfying

$$\langle z|z'\rangle := \langle z||z'\rangle = K(z,z') \quad \text{for } z,z' \in Z,$$

with the linear functionals[3]

$$\langle z| := |z\rangle^*$$

acting on $\mathbb{Q}^\times(Z)$. Thus, there is a distinguished map from Z to $\mathbb{Q}(Z)$ mapping z to the vectors $|z\rangle$ ($z \in Z$); these are called the **coherent states** of Z in $\mathbb{Q}(Z)$. In this book, we use this Dirac bra/ket notation *only* for coherent states and their adjoints.

We call the completion $\overline{\mathbb{Q}}(Z) := \overline{\mathbb{Q}(Z)}$ of a quantum space the corresponding **completed quantum space** of Z. The corresponding **augmented quantum space** is the antidual $\mathbb{Q}^\times(Z) := \mathbb{Q}(Z)^\times$. We have

$$\mathbb{Q}(Z) \subseteq \overline{\mathbb{Q}}(Z) \subseteq \mathbb{Q}^\times(Z).$$

If the quantum space is infinite-dimensional, then $\mathbb{Q}(Z)$ is usually a proper subspace of the Hilbert space $\overline{\mathbb{Q}}(Z)$. By definition of the weak-$\star$ topology of $\mathbb{Q}^\times(Z)$, $\psi_\ell \in \mathbb{Q}^\times(Z)$ converges to $\psi \in \mathbb{Q}^\times(Z)$ iff $\langle z|\psi_\ell \to \langle z|\psi$ for all $z \in Z$.

Proposition 5.3.1.

(i) *Let $\mathbb{H}$ be a Euclidean space. Then for any set Z and any mapping $c : Z \to \mathbb{H}$,*

$$K(z,z') := c(z)^* c(z') \tag{5.3}$$

defines a coherent product on Z that turns Z into a coherent space, whose quantum space $\mathbb{Q}(Z)$ is the space consisting of the finite linear combinations of coherent states $|z\rangle := c(z)$. ($\mathbb{Q}(Z)$ is usually a proper subspace of $\mathbb{H}$.)

3 With this convention, $\langle z|$ is a linear functional mapping $\psi \in \mathbb{Q}^\times(Z)$ to $\langle z|\psi$, whereas $|z\rangle \in \mathbb{Q}(Z)$ is an antilinear functional mapping $\psi \in \mathbb{Q}^\times(Z)$ to $\psi^*|z\rangle$.

(ii) *Conversely, every coherent product can be written in the form* (5.3) *such that the coherent states are given as* $|z\rangle = c(z)$.

Proof. (i) follows by combining Example 5.2.1(i) with the definition of the quantum space. To see (ii), take $\mathbb{H} = \mathbb{Q}(Z)$ and define $c(z) := |z\rangle$. □

Theorem 5.3.2. *Every coherent space Z has a quantum space $\mathbb{Q}(Z)$. It is unique up to an isomorphism that maps coherent states with the same label to each other.*

Proof. By definition of a coherent space, the coherent product K is of positive type. Hence, the Moore–Aronszejn theorem (Theorem 4.6.1) applies and provides a Hilbert space $\overline{\mathbb{Q}}$. If we define the coherent states $|z\rangle := q_z$ and their adjoints $\langle z| := q_z^*$, we find from (4.15) below that

$$\langle z|z'\rangle = \langle q_z, q_{z'}\rangle = K(z, z').$$

Thus, the space $\mathbb{Q}$ consisting of the finite linear combinations of coherent states is a quantum space. If $\mathbb{Q}$ and $\mathbb{Q}'$ are quantum spaces for Z with coherent states $|z\rangle$ and $|z\rangle'$, respectively, then

$$I(\phi) := \sum a_k |z_k\rangle' \quad \text{if } \phi = \sum a_k |z_k\rangle$$

defines a map $I : \mathbb{Q} \to \mathbb{Q}'$. Indeed, if $\phi = \sum b_k |z_k\rangle$ is another representation of ϕ, then

$$\sum a_k K(z, z_k) = \sum a_k \langle z|z_k\rangle = \langle z|\phi = \sum b_k \langle z|z_k\rangle = \sum b_k K(z, z_k).$$

Therefore, $\phi' := \sum b_k |z_k\rangle'$ satisfies

$$\begin{aligned} {}'\langle z|\phi' &= \sum b_k {}'\langle z|z_k\rangle' = \sum b_k K(z, z_k) \\ &= \sum a_k K(z, z_k) = \sum a_k {}'\langle z|z_k\rangle' = {}'\langle z|I(\phi) \end{aligned}$$

for all $z \in Z$, whence $\phi' = I(\phi)$. This map is easily seen to be an isomorphism. □

Note that any linear map from a quantum space of a coherent space into $\mathbb{C}$ is continuous, and any linear map from a quantum space of a coherent space into its antidual is also continuous.

Let Z, Z' be coherent spaces. A **morphism** from Z to Z' is a map $\rho : Z \to Z'$ such that

$$K'(\rho(z), \rho(w)) = K(z, w) \quad \text{for } z, w \in Z; \tag{5.4}$$

if $Z' = Z$, then ρ is called an **endomorphism**. Two coherent spaces Z and Z' with coherent products K and K', respectively, are called **isomorphic** if there is a bijective morphism $\rho : Z \to Z'$. In this case, we write $Z \cong Z'$, and we call the map $\rho : Z \to Z'$

an **isomorphism** of the coherent spaces. Clearly, $\rho^{-1} : Z' \to Z$ is then also an isomorphism. If $Z' = Z$ and $K' = K$, then we call ρ an **automorphism** of Z. Automorphisms are closely related to the more general concept of coherent maps, introduced in Neumaier & Ghaani Farashahi [212].

Proposition 5.3.3. *Let Z, Z' be coherent spaces, and let $\rho : Z \to Z'$ be an isomorphism. Then*
(i) $K(\rho^{-1}(z'), \rho^{-1}(w')) = K'(z', w')$ *for all* $z', w' \in Z'$.
(ii) $K'(z', \rho(z)) = K(\rho^{-1}(z'), z)$ *for all* $z \in Z$ *and* $z' \in Z'$.

Proof. (i) is straightforward.
(ii) Let $z \in Z$ and $z' \in Z'$. Then

$$K'(z', \rho(z)) = K'(\rho(\rho^{-1}(z')), \rho(z)) = K(\rho^{-1}(z'), z). \qquad \square$$

Proposition 5.3.4. *Let Z be a coherent space with coherent product K, and let Z' be an arbitrary set. Then for any map $\rho : Z' \to Z$,*

$$K'(z, z') := K(\rho z, \rho z') \quad \text{for } z, z' \in Z'$$

defines a coherent product on Z'. This turns Z' into a coherent space, with respect to which ρ is a morphism.

Proposition 5.3.5. *Let Z, Z' be isomorphic coherent spaces. Then any two quantum spaces $\mathbb{Q}(Z)$ of Z and $\mathbb{Q}(Z')$ of Z' are isometric Euclidean spaces.*

Proof. Let Z, Z' be isomorphic coherent spaces. Let $\mathbb{Q}(Z)$ and $\mathbb{Q}(Z')$ be quantum spaces of Z and Z', respectively. Let $\rho : Z \to Z'$ be an isomorphism of coherent spaces. We define the map $T_\rho : \mathbb{Q}(Z) \to \mathbb{Q}(Z')$ by

$$T_\rho\Big(\sum_k c_k |z_k\rangle\Big) := \sum_k c_k |\rho(z_k)\rangle' \quad \text{for all } \sum_k c_k |z_k\rangle \in \mathbb{Q}(Z).$$

Now

$$\begin{aligned}\Big\| T_\rho\Big(\sum_k c_k |z_k\rangle\Big)\Big\|^2_{\mathbb{Q}(Z')} &= \Big\| \sum_k c_k |\rho(z_k)\rangle' \Big\|^2_{\mathbb{Q}(Z')} \\ &= \sum_k \sum_j \overline{c_k} c_j K'(\rho(z_k), \rho(z_j)) \\ &= \sum_k \sum_j \overline{c_k} c_j K(z_k, z_j) = \Big\| \sum_k c_k |z_k\rangle \Big\|^2_{\mathbb{Q}(Z)}.\end{aligned}$$

This implies that T_ρ is a well-defined isometry. Since ρ is surjective, T_ρ is surjective as well. Thus, T_ρ is an isomorphism. $\square$

5.4 Length, angle, distance

The definition of a coherent space has a number of simple but useful general consequences. The Hermiticity of the Gram matrix of z, z' gives

$$\overline{K(z,z')} = K(z',z). \tag{5.5}$$

Since the diagonal elements of a Hermitian positive semidefinite matrix are real and nonnegative:

$$K(z,z) \geq 0 \quad \text{for all } z \in Z. \tag{5.6}$$

In particular, we may define the **length** of $z \in Z$ to be

$$n(z) := \sqrt{K(z,z)} \geq 0. \tag{5.7}$$

Since every principal submatrix of a Hermitian positive semidefinite matrix has real nonnegative determinants, the determinants of size 2 lead to

$$\left|K(z,z')\right|^2 \leq K(z,z)K(z',z'). \tag{5.8}$$

Taking square roots gives the **coherent Cauchy–Schwarz inequality**

$$\left|K(z,z')\right| \leq n(z)n(z'). \tag{5.9}$$

This allows us to define the **angle** between two points $z, z' \in Z$ of positive length by

$$\angle(z,z') := \arccos \frac{|K(z,z')|}{n(z)n(z')} \in [0,\pi[. \tag{5.10}$$

Theorem 5.3.2 implies that, in a sense, coherent spaces are just the subsets of Euclidean spaces. However, separating the structure of a coherent space Z from the notion of a Euclidean space allows many geometric features to be expressed in terms of Z and the coherent product alone, without direct references to the quantum space. The latter only serves as a convenient tool for proving assertions of interest. For example, the study of symmetry in Neumaier & Ghaani Farashahi [212] benefits from this separation. Another example is the distance function induced on Z by the Euclidean distance, as in the proof of Proposition 5.4.1 (below).

Proposition 5.4.1 (Parthasarathy & Schmidt [228, Corollary 1.3/4]). *The **distance***

$$d(z,z') := \sqrt{K(z,z) + K(z',z') - 2\operatorname{Re} K(z,z')} \tag{5.11}$$

of two points $z, z' \in Z$ is nonnegative and satisfies the triangle inequality. With (5.7)*, we have*

$$\left|n(z) - n(z')\right| \leq d(z,z') \leq n(z) + n(z'), \tag{5.12}$$

$$\left|K(y,z) - K(y',z')\right| \leq d(y,y')n(z') + n(y)d(z,z'). \tag{5.13}$$

Proof. The expression under the square root of (5.11) is

$$\langle z|z\rangle + \langle z'|z'\rangle - \langle z|z'\rangle - \langle z'|z\rangle = \big\||z\rangle - |z'\rangle\big\|^2, \tag{5.14}$$

whence $d(z,z')$ is the Euclidean distance between $|z\rangle$ and $|z'\rangle$. This implies nonnegativity and the triangle inequality. $n(z)$ is the length of $|z\rangle$, and (5.12) follows. The Cauchy–Schwarz inequality gives

$$\big|K(y,z) - K(y,z')\big| = \big|\langle y|(|z\rangle - |z'\rangle)\big| \le n(y)d(z,z').$$

Hence,

$$\begin{aligned}\big|K(y,z) - K(y',z')\big| &= \big|K(y,z) - K(y,z') + K(y,z') - K(y',z')\big| \\ &\le \big|K(y,z') - K(y',z')\big| + \big|K(y,z) - K(y,z')\big| \\ &\le d(y,y')n(z') + d(z,z')n(y).\end{aligned}$$

This proves (5.13). □

We call a coherent space **nondegenerate** if $K(z'',z') = K(z,z')$ for all $z' \in Z$ implies $z'' = z$. Clearly, this is the case iff the mapping from Z to $\mathbb{Q}(Z)$ that maps each $z \in Z$ to the corresponding coherent state $|z\rangle$ is injective.

Proposition 5.4.2. *The distance map d is a metric on Z iff K is nondegenerate on Z.*

Proof. (5.14) implies that $d(z,z') = 0$ iff $|z\rangle = |z'\rangle$. Hence, d is a metric on Z iff K is nondegenerate on Z. □

The distance map d is a quasimetric. Therefore, it induces in the standard way a topology on Z called the **metric topology**, and denoted by τ_m. There is a second topology on nondegenerate coherent spaces Z, the **coherent topology** denoted by τ_c, defined by calling a net z_ℓ **coherently convergent** to z iff $K(z_\ell,z') \to K(z,z')$ for all $z' \in Z$. It can be readily checked that the coherent topology τ_c is at least as fine as the metric topology τ_m, because if $z_n \to z$ in the metric topology, then also $z_n \to z$ in the coherent topology.

Theorem 5.4.3. *In any coherent space, the metric topology is the weakest (coarsest) topology in which K is continuous.*

Proof. We equip $Z \times Z$ with the product topology induced by the metric topology on Z. Let (z_n, z_n') be a convergent sequence to $(z,z') \in Z \times Z$. Then $z_n \to z$ and $z_n' \to z'$ in the metric topology. Hence, the sequence of $n(z_n)$ is bounded. Thus,

$$\big|K(z_n,z_n') - K(z,z')\big| \le d(z_n,z)n(z') + d(z_n',z')n(z_n),$$

which implies that $\lim_n K(z_n,z_n') = K(z,z')$.

Now let τ be any topology on the coherent space Z such that $K : Z \times Z \to \mathbb{C}$ is continuous. To prove that τ is at least as fine as τ_m, we assume that $w_n \to w$ in Z with respect to τ. Since K is continuous with respect to τ, and $K(w,w) = n(w)^2$, we find

$$\begin{aligned}\lim_n d(w_n, w) &= \lim_n \sqrt{K(w_n, w_n) + n(w)^2 - 2\operatorname{Re} K(w_n, w)} \\ &= \sqrt{K(w,w) + n(w)^2 - 2\operatorname{Re} K(w,w)} \\ &= \sqrt{n(w)^2 + n(w)^2 - 2\operatorname{Re} n(w)^2} = 0,\end{aligned}$$

which implies that $w_n \to w$ in Z with respect to the metric topology as well. Thus, τ is at least as fine as τ_m. This implies that the metric topology is the weakest (coarsest) topology, in which K is continuous. □

A **coherent manifold** is a smooth (= C^∞) real manifold Z with a smooth coherent product $K : Z \times Z \to \mathbb{C}$, with which Z is a coherent space. In a nondegenerate coherent manifold, the infinitesimal distance equips the manifold with a canonical Riemannian metric. Coherent manifolds are studied in detail in Neumaier & Ghaani Farashahi [213].

5.5 Vectors in the augmented quantum space

By definition, all vectors in $\mathbb{Q}^\times(Z)$ can be constructed as weak-* limits of nets in $\mathbb{Q}(Z)$, and hence by the following construction:

Proposition 5.5.1. *A net ψ_ℓ in $\mathbb{Q}(Z)$ is weak-* convergent iff, for all $z \in Z$, the inner product $\langle z|\psi_\ell$ converges. In this case, the limit $\psi = \lim \psi_\ell \in \mathbb{Q}^\times(Z)$ is characterized by*

$$\psi(|z\rangle) := \langle z|\psi = \lim \langle z|\psi_\ell \quad \text{for } z \in Z. \tag{5.15}$$

Proof. By definition, weak-* convergence to ψ says that $\psi_\ell(\phi) \to \psi(\phi)$ for all $\phi \in \mathbb{Q}(Z)$. In particular, $\langle z|\psi_\ell = \psi_\ell(|z\rangle)$ converges to $\psi(|z\rangle)$, and (5.15) holds. Conversely, suppose that the limit (5.15) exists for all $z \in Z$. Any $\phi \in \mathbb{Q}(Z)$ can be written as a finite linear combination $\phi = \sum_k \alpha_k |z_k\rangle$. Hence, $\psi_\ell(\phi) = \sum_k \alpha_k \psi_\ell(|z_k\rangle) = \sum_k \langle z_k|\psi_\ell$ converges. Thus, the net ψ_ℓ is weak-* convergent. □

Let X be an open subset of a finite-dimensional complex vector space. We call a map $u : X \to Z$ **smooth** if, for each $z \in Z$, $K(z, u(x))$ is C^∞ as a function of $x \in X$, and **strongly smooth** if, in addition, $K(u(x), u(y))$ is C^∞ in $(x,y) \in X \times X$. (With these notions of smoothness, Z becomes in two ways a diffeological space; see Iglesias-Zemmour [143].)

Theorem 5.5.2. *Let $u : X \to Z$ be a smooth map, and let A be a linear differential operator on $C^\infty(X)$. Then the following hold:*

(i) *For any $x \in X$, there is a unique state $\psi_{u,A,x} \in \mathbb{Q}^\times(Z)$ such that*

$$\langle z|\psi_{u,A,x} = A(x)K(z,u(x)). \tag{5.16}$$

(ii) *For fixed u, x, the map $A \to \psi_{u,A,x}$ is linear.*
(iii) *If u is strongly smooth, then $\psi_{u,A,x} \in \overline{\mathbb{Q}}(Z)$.*
(iv) *If $u : X \to Z$ and $v : Y \to Z$ are strongly smooth, if A, B are linear differential operators on $C^\infty(X)$ and $C^\infty(Y)$, respectively, and $x \in X$, $y \in Y$, then*

$$\psi^*_{u,A,x}\psi_{v,B,y} = \overline{A(x)\overline{B(y)K(u(x),u(y))}}.$$

Proof. (i) Every linear differential operator A on $C^\infty(X)$ can be written as a limit of a sequence of finite linear combinations of function values,

$$A(x)f(x) = \lim_{\ell\to\infty} \sum_k \alpha_{\ell k} f(x + h_{\ell k}),$$

obtainable by replacing each derivative by a limit of difference quotients. We apply this to the function $f \in C^\infty(X)$, defined by $f(x) := K(z,u(x))$, and find that the sequence of vectors $\psi_\ell := \sum_k \alpha_{\ell k}|u(x + h_{\ell k})\rangle$ satisfies

$$\begin{aligned}\lim_\ell \langle z|\psi_\ell &= \lim_\ell \sum_k \alpha_{\ell k}\langle z|u(x + h_{\ell k})\rangle = \lim_\ell \sum_k \alpha_{\ell k} K(z,u(x + h_{\ell k}))\\ &= \lim_\ell A(x)K(z,u(x)).\end{aligned}$$

Thus, $\psi := \lim \psi_\ell$ exists and satisfies (5.16).

(ii) is straightforward.

(iii) and (iv) are proved similarly. □

The above expressions for inner products make many calculations easy that are difficult in Hilbert spaces, whose inner product is defined through a measure. In particular, the quantum spaces of coherent spaces with an easily computable coherent product can be studied efficiently without measures or integrations, in terms of the explicit coherent product and differentiation only.

5.6 New states from old ones

From the set of coherent states, it is possible to create a large number of other states, whose inner product is computable by a closed formula. This is important for numerical applications, since one can pick from the new states created in this fashion a suitable subset, and declare the states belonging to this subset to be the coherent states of a new, derived coherent space. This way of constructing new coherent spaces from old ones allows one to apply the general body of techniques for the analysis of coherent spaces and their quantum properties to the new coherent space. In this way, many

known numerical techniques for quantum physics problems become organized in the same setting.

The first, often useful construction takes a path $u(t)$ in Z and creates new states

$$[R_t u(t)] := \lim_{h \downarrow 0} h^{-1}(|u(t+h)\rangle - |u(t)\rangle).$$

We write $= \partial_j K$ for the partial derivative with respect to the jth argument of K and find the inner products

$$\begin{aligned}
\langle z|[R_t u(t)] &= \lim_{h \downarrow 0} h^{-1}(K(z, u(t+h)) - K(z, u(t))) = \partial_2 K(z, u(t))\dot{u}(t), \\
[R_t u(t)]^* [R_s v(s)] &= \lim_{h \downarrow 0} h^{-1}(\langle u(t+h)|[R_s v(s)] - \langle u(t)|[R_s v(s)]) \\
&= \lim_{h \downarrow 0} h^{-1}(\partial_2 K(u(t+h), v(s))\overline{\dot{v}(t)} - \partial_2 K(u(t), v(s))\overline{\dot{v}(t)}) \\
&= \dot{u}(t)\partial_1\partial_2 K(u(t), v(s))\dot{v}(s).
\end{aligned}$$

Similar expressions can be found by taking other smooth parameterizations of submanifolds of Z and taking the limits corresponding to first-order or higher-order partial derivatives.

A trivial construction is to take linear combinations

$$[\alpha, y] := \sum_k \alpha_k |y_k\rangle,$$

where α is a finite sequence of complex numbers α_k, and y is a finite sequence of points $y_k \in Z$. The inner products are given by

$$\begin{aligned}
\langle z|[\alpha, y] &= \sum_k \alpha_k K(z, y_k), \\
[\alpha, y]^* [\alpha', y]' &= \sum_{j,k} \alpha_j \alpha'_k K(y_j, y_k).
\end{aligned}$$

This also works for infinite sequences, provided the right-hand sides are always absolutely convergent, and with sums replaced by integrals for weighted integrals $\int \alpha(x)|y(x)\rangle\, d\mu(x)$, provided the corresponding integrals on the right-hand sides are always absolutely convergent. Of course, all these recipes can also be combined.

We see that, unlike in traditional Hilbert spaces, where the calculation of inner products always requires to evaluate often high-dimensional integrals, here the calculation of inner products is much simpler, often only taking sums and derivatives.

5.7 Some examples

We now give a long list of basic examples of coherent spaces exhibiting the flavor of the relations to other fields of mathematics and science. As indicated in the introduction, this is just the tip of an iceberg; many other coherent spaces will be discussed elsewhere.

The first group of examples arises in applications to quantum mechanics. For the physical background see, for example, Neumaier & Westra [214].

Examples 5.7.1. The simplest instances of coherent spaces are the spaces formed by the subsets Z of $\mathbb{C}^n$, which are closed under conjugation and scalar multiplication, with one of the coherent products:

$$K(z,z') := \begin{cases} 1 & \text{if } z' = \overline{z}, \\ 0 & \text{otherwise,} \end{cases} \tag{5.17}$$

$$K(z,z') := z^* z', \tag{5.18}$$

$$K(z,z') := (z^* z')^{2j} \quad \left(j = 0,\ \frac{1}{2},\ 1,\ \frac{3}{2},\ \dots\right), \tag{5.19}$$

$$K(z,z') := e^{(z^* z' - \frac{1}{2}\|z\|^2 - \frac{1}{2}\|z'\|^2)/\hbar}, \tag{5.20}$$

where $\hbar$ is a positive real number. In the applications to quantum mechanics, $\hbar$ is the **Planck constant**. The axioms are easily verified using the constructions of Proposition 4.5.2 and Theorem 4.5.4.

(i) $Z = \mathbb{C}^n$ with the coherent product (5.17) corresponds to the phase space of a classical system of n oscillators, with n position and n momentum degrees of freedom, via the identification

$$z = q + ip, \quad q = \operatorname{Re} z, \quad p = \operatorname{Im} z.$$

In the corresponding quantum space, the associated coherent states are orthonormal basis vectors, indexed by the phase space points.

(ii) The unit sphere Z in $\mathbb{C}^2$ with the coherent product (5.17) corresponds to the phase space of a **classical spin**, such as a polarized light beam or a spinning top fixed at its point.

(iii) The unit ball Z in $\mathbb{C}^2$ with the coherent product (5.17) corresponds to the classical phase space of (monochromatic) **partially polarized light**.

(iv) $Z = \mathbb{C}^n$ with the coherent product (5.18) has as quantum space the Hilbert space $\mathbb{C}^n$ of an n-level quantum system. The associated coherent states are all state vectors.

(v) The unit sphere

$$Z = \{z \in \mathbb{C}^2 \mid z^* z = 1\}$$

in $\mathbb{C}^2$ with the coherent product (5.19) corresponds to the **Bloch sphere** representing a single quantum mode of an atom with angular momentum (**spin**) j. The corresponding quantum space has dimension $2j + 1$. For $j = 1/2$, it also represents an arbitrary 2-level system, a so-called **qubit**; the prime example is the polarization of a single photon mode (then Z is called the **Poincaré sphere**). The associated coherent states are the so-called **spin coherent states**. This example shows that

a given set Z may carry more than one interesting coherent product, resulting in different coherent spaces with nonisomorphic quantum spaces. For $j \to \infty$, these spaces degenerate into the coherent space of a classical spin.

(vi) $Z = \mathbb{C}^n$ with the coherent product (5.20) has as quantum space, the bosonic Fock space, with n degrees of freedom, corresponding to n independent harmonic oscillators. The associated coherent states are the so-called **Glauber coherent states**. In the so-called classical limit $\hbar \to 0$ (which can be taken mathematically, though not in reality), the space degenerates into the coherent space of a classical system with n spatial degrees of freedom.

We note that for (5.20), the power construction from Theorem 4.5.4(i) amounts to a replacement of $\hbar$ by $\hbar/n$. Therefore, the classical limit amounts here to applying the power construction for arbitrary n and considering the limit $n \to \infty$. Generalizing this to arbitrary coherent spaces provides a general definition of the **classical limit**, even when $\hbar$ does not appear in the coherent product. For example, the power construction applied to (5.18) produces (5.19) with $2j = n$. Thus, the classical limit amounts here to the limit of infinite angular momentum. The classical limit and a related semiclassical analysis of coherent quantum physics is investigated in general in Neumaier [209].

It is interesting to note that in order that a coherent product results, $\hbar$ can take in (5.20) any positive value, whereas in (5.19), $2j$ must be a nonnegative integer. The latter is already needed in order that the power is unambiguously defined.

More generally, any coherent space Z gives rise to an infinite family of coherent spaces Z_n on the same set Z, but with modified coherent product $K_n(z, z') := K(z, z')^n$ with a nonnegative integer n. The quantum space $\mathbb{Q}(Z_n)$ is the symmetric tensor product of n copies of the quantum space $\mathbb{Q}(Z)$.

In general, the need for a nonnegative integer in the exponent is related to **Bohr–Sommerfeld quantization** of compact phase spaces. The conditions that must be imposed on the exponent in general are captured through the concept of a Berezin–Wallach set in Section 4.10.

Example 5.7.2. The set $Z = \mathbb{R}_+$ of positive real numbers is a real coherent space with trivial conjugation for any of the coherent products:

$$K(z, z') = \min(z, z'),$$
$$K(z, z') = (z + z')^{-1}.$$

(i) In the first case, a completed quantum space is $L^2(\mathbb{R}_+)$ with coherent states

$$k_z(z') = \begin{cases} 1 & \text{if } z' \le z, \\ 0 & \text{otherwise.} \end{cases}$$

(ii) In the second case, a completed quantum space is $L^2(\mathbb{R}_+)$ with coherent states

$$k_z(z') = e^{-zz'}$$

since

$$\langle k_z, k_{z'} \rangle = \int_{0+}^{\infty} dy \overline{k_z(y)} k_{z'}(y) = \int_0^{\infty} dy e^{-zy} e^{-z'y} = \int_0^{\infty} dy e^{-(z+z')y} = \frac{1}{z+z'}.$$

The following spaces are important not only in complex analysis, but are also relevant in quantum physics, for the analysis of quantum mechanical scattering problems (DE BRANGES & ROVNYAK [51, Theorem 4]). Example 5.7.3(ii) (below) is relevant in signal processing.

For a function $f : Z \subset \mathbb{C} \to \mathbb{C}$, we define its conjugate $\overline{f} : Z \to \mathbb{C}$ by

$$\overline{f}(z) := \overline{f(\overline{z})}. \tag{5.21}$$

Examples 5.7.3.

(i) A **de Branges function** is an entire analytic function $E : \mathbb{C} \to \mathbb{C}$ satisfying

$$|E(\overline{z})| < |E(z)| \quad \text{if } \operatorname{Im} z > 0. \tag{5.22}$$

With the coherent product

$$K(z, z') := \begin{cases} \overline{E}'(\overline{z})E(z') - E'(\overline{z})\overline{E}(z') & \text{if } z' = \overline{z}, \\ \frac{\overline{E}(\overline{z})E(z') - E(\overline{z})\overline{E}(z')}{2i(\overline{z} - z')} & \text{otherwise,} \end{cases}$$

$Z = \mathbb{C}$ is a coherent space. A corresponding quantum space is the subspace of $L^2(\mathbb{R})$, spanned by the coherent states q_z, denoted by $\mathcal{H}(E)$, and defined by

$$q_z(t) = \frac{K(\overline{z}, t)}{E(t)} := \lim_{\varepsilon \downarrow 0} \frac{K(\overline{z}, t + i\varepsilon)}{E(t + i\varepsilon)} \quad \text{for } t \in \mathbb{R}.$$

(The denominator on the right is nonzero by (5.22). The limit exists and is continuous as a function of t since at an n-fold zero t of E, the function $K(\overline{z}, \cdot)$ has t as a zero of multiplicity at least r.) Indeed, the formula $q_z^* q_{z'} = K(z, z')$ follows by evaluating the integral expression for $q_z^* q_{z'}$ using the residue theorem. For details see DE BRANGES [50, Theorem 19, p. 50], where the quantum space is more fully characterized.

(ii) $Z = \mathbb{C}$ is a coherent space with the coherent product

$$K(z, z') := \operatorname{sinc}(\overline{z} - z'), \quad \operatorname{sinc}(z) := \begin{cases} 1 & \text{if } z = 0, \\ \sin(z)/z & \text{otherwise.} \end{cases}$$

This is the special case $E(z) = e^{-iz}$ of (i).

(iii) A **Schur function** (Schur [272]) is an analytic function s from the open unit disk $B(0;1)$ in $\mathbb{C}$ into its closure. With the coherent product

$$K(z,z') := \frac{1-\overline{s(z)}s(z')}{1-\bar{z}z'},$$

$Z := B(0;1)$ is a coherent space. Note that the inverse is defined since $|\bar{z}z'| < 1$. Coherence follows from results by de Branges & Rovnyak [52]. The corresponding quantum spaces are the sub-Hardy spaces discussed by Sarason [260], also called **de Branges–Rovnyak spaces**; see the recent survey by Ball & Bolotnikov [21].

(iv) The **Szegö space** is the coherent space defined on the open unit disk in $\mathbb{C}$,

$$D(0,1) := \{z \in \mathbb{C} \mid |z| < 1\},$$

by the coherent product

$$K(z,z') := (1-\bar{z}z')^{-1},$$

see Szegö [282]. This example from 1911 is probably the earliest nontrivial explicit, coherent product in the literature. It is the special case $s = 0$ of (iii); its quantum space is the Hardy space on the unit disk. Coherence also follows directly from Theorem 4.5.4(iii).

In general, unlike in these (and other simple) examples, there need not be a simple realization of a quantum space in terms of an L^2 space with respect to a suitable measure. Fortunately, such a description is usually not needed in applications to physics since one can work comfortably in the quantum space using only its defining properties. This is one of the strengths of the concept of coherent spaces, as it allows one to avoid the often cumbersome evaluation of integrals in the computation of inner products.

5.8 Normal, projective, and nondegenerate coherent spaces

We call a coherent space **normal** if

$$\begin{cases} K(z,z') = 1 & \text{if } z' = z, \\ |K(z,z')| < 1 & \text{otherwise.} \end{cases}$$

In a normal coherent space, coherent states have norm 1. Hence, the distance simplifies to

$$d(z,z') := c\sqrt{1-\operatorname{Re}K(z,z')}, \quad c = \sqrt{2}. \tag{5.23}$$

This distance was studied by Arcozzi et al. [14] with $c = 1$ rather than the above value. (5.23) implies that a normal coherent space Z is nondegenerate.

Proposition 5.8.1. *Let Z be a coherent space with coherent product K. Then, for any function $\gamma : Z \to \mathbb{C}$, the set Z with* ***scaled coherent product***

$$K_\gamma(z, z') := \overline{\gamma(z)} K(z, z') \gamma(z')$$

is also a coherent space.

Proof. The Gram matrix G' of the scaled coherent product has entries

$$G_{jk} := K_\gamma(z_j, z_k) = \overline{\gamma(z_j)} K(z_j, z_k) \gamma(z_k)$$

and is clearly Hermitian. For any vector u, we define the vector v with components $v_k := \gamma(z_k) u_k$ and find

$$u^* G' u = \sum_{j,k} \overline{u}_j \overline{\gamma(z_j)} K(z_j, z_k) \gamma(z_k) u_k = \sum_{j,k} \overline{v}_j K(z_j, z_k) v_k \geq 0.$$

Thus, G' is positive semidefinite. □

Proposition 5.8.2. *Let Z be a coherent space. If the coherent product is not identically zero, then there is a normal, coherent space Z' such that there is an isomorphism $\alpha : \mathbb{Q}(Z) \to \mathbb{Q}(Z')$ with*

$$\{\alpha|z\rangle \mid z \in Z\} \subseteq \{\lambda|z'\rangle \mid \lambda \in \mathbb{C},\ z' \in Z'\}.$$

Therefore, any image of a coherent state of Z is a multiple of some coherent state of Z'.

Proof. If $K(z, z) = 0$, then the coherent state $|z\rangle$ vanishes by the Cauchy–Schwarz inequality (4.4). Thus, we can delete such points from Z. By scaling using Proposition 5.8.1, we may assume that $K(z, z) = 1$ without changing the Hilbert space. Now the proof of the Cauchy–Schwarz inequality (4.4) shows that if $|K(z, z')| = 1$, then the coherent states $|z\rangle$ and $|z'\rangle$ differ by a phase only; so we may delete one of them without changing the Hilbert space. The new coherent space is normal. □

We call a coherent space Z **projective** if there is a **scalar multiplication** that assigns to each $\lambda \in \mathbb{C}^\times$, and each $z \in Z$ a point $\lambda z \in Z$ such that

$$K(z, \lambda z') = \lambda^e K(z, z') \quad \text{for all } z, z' \in Z, \tag{5.24}$$

for some $e \in \mathbb{Z} \setminus \{0\}$ called the **degree**. Note that a coherent space cannot be both normal and projective. Example 5.7.1(v) is projective of degree $e = 2j$; Examples 5.7.2(i) and (ii) are projective of degree 1 and –1, respectively.

There are important degenerate projective spaces, where the scalar multiplication is not associative, because it is not canonically defined. An example are the Klauder spaces from Example 5.2.2, which are projective of degree 1 with the scalar multiplication

$$\lambda[z_0, \mathbf{z}] := [z_0 + \log \lambda, \mathbf{z}],$$

using an arbitrary but fixed branch of log. The need to restrict to a fixed branch causes the associative law to be not valid universally. On the other hand, we have:

Proposition 5.8.3. *Let Z be a nondegenerate and projective space. Then the scalar multiplication is associative:*

$$\lambda(\mu z) = (\lambda\mu)z \quad \text{for } \lambda, \mu \in \mathbb{C}^\times,\ z \in Z. \tag{5.25}$$

Proof. Let $z \in Z$ and $\lambda, \mu \in \mathbb{C}^\times$. For all $z' \in Z$, we have

$$K(\lambda(\mu z), z') = \overline{\lambda}^e K(\mu z, z') = \overline{\lambda}^e \overline{\mu}^e K(z, z') = (\overline{\lambda\mu})^e K(z, z') = K((\lambda\mu)z, z').$$

Now nondegeneracy of K implies (5.25). □

Proposition 5.8.4. *Let Z be a projective coherent space of degree e. Then*

$$K(\lambda z, z') = \overline{\lambda}^e K(z, z') \quad \text{for all } z, z' \in Z, \tag{5.26}$$

$$|\lambda z\rangle = \lambda^e |z\rangle \quad \text{for } \lambda \in \mathbb{C}^\times,\ z \in Z, \tag{5.27}$$

$$K(z, \lambda z') = K(\overline{\lambda} z, z') \quad \text{for } \lambda \in \mathbb{C}^\times,\ z \in Z. \tag{5.28}$$

Proof. (5.26) follows from the definition and (5.5). To prove (5.27), let $z \in Z$ and $\lambda \in \mathbb{C}$. Then, for all $z' \in Z$,

$$\langle z' | \lambda z \rangle = K(z', \lambda z) = \lambda^e K(z', z) = \lambda^e \langle z' | z \rangle.$$

Finally, using (5.24) and (5.26), we get

$$K(z, \lambda z') = \lambda^e K(z, z') = K(\overline{\lambda} z, z').$$ □

Formula (5.28) implies that the scalar multiplication maps are coherent maps in the sense introduced in Section 5.9 below.

Any coherent space can be extended to a projective coherent space without changing the quantum space. The idea of a projective extension can be traced back to Klauder [158].

Proposition 5.8.5. *Let Z be a coherent space and e be a nonzero integer. Then the* ***projective extension*** *$PZ := \mathbb{C}^\times \times Z$ of degree e is a projective coherent space with coherent product*

$$K_{\mathrm{pe}}((\lambda, z), (\lambda', z')) := \overline{\lambda}^e K(z, z') \lambda'^e \tag{5.29}$$

and scalar multiplication $\lambda'(\lambda, z) := (\lambda'\lambda, z)$. The corresponding quantum spaces $\mathbb{Q}(Z)$ and $\mathbb{Q}(PZ)$ are canonically isomorphic.

Proof. It is straightforward to check that PZ, with respect to the projective extension kernel K_{pe}, is a projective coherent space of degree e. The map $T : \mathbb{Q}(PZ) \to \mathbb{Q}(Z)$ given by

$$T\sum_\ell c_\ell|(\lambda_\ell,z_\ell)\rangle := \sum_\ell c_\ell\lambda_\ell^e|z_\ell\rangle \quad \text{for all } \sum_\ell c_\ell|(\lambda_\ell,z_\ell)\rangle \in \mathbb{Q}(PZ)$$

is well-defined and linear. Also, we have

$$\begin{aligned}\Big\|T\sum_\ell c_\ell|(\lambda_\ell,z_\ell)\rangle\Big\|^2_{\mathbb{Q}(Z)} &= \Big\|\sum_\ell c_\ell\lambda_\ell^e|z_\ell\rangle\Big\|^2_{\mathbb{Q}(Z)}\\ &= \sum_j\sum_k \overline{c_j}\overline{\lambda_j^e}c_k\lambda_k^e K(z_j,z_k)\\ &= \sum_j\sum_k \overline{c_j}c_k K_{\mathrm{pe}}((\lambda_j,z_j),(\lambda_k,z_k))\\ &= \Big\|\sum_\ell c_\ell|(\lambda_\ell,z_\ell)\rangle\Big\|^2_{\mathbb{Q}(PZ)},\end{aligned}$$

which implies that T is an isometric linear operator. Thus, T is injective as well. Let $\psi = \sum_\ell c_\ell|z_\ell\rangle \in \mathbb{Q}(Z)$ with $c_\ell \neq 0$ for all ℓ. Then $\phi := \sum_\ell |(c_\ell^{-e},z_\ell)\rangle \in \mathbb{Q}(PZ)$ with $T\phi = \psi$. Thus, T is an isomorphism. □

Corollary 5.8.6. *Let Z be a coherent space, and let e be a nonzero integer. Then Z is projective of degree e iff $P_eZ \cong Z$.*

Proof. Let Z be a projective space of degree $e \in \mathbb{Z}$. We then define $\rho : P_eZ \to Z$ by $\rho(\lambda,z) := \lambda z$ for all $(\lambda,z) \in P_eZ$. It is easy to check that $\rho : P_eZ \to Z$ is an isomorphism. Hence, $P_eZ \cong Z$. Conversely, suppose that $P_eZ \cong Z$, and let $\rho : P_eZ \to Z$ be an isomorphism of coherent spaces. Then, with multiplication defined by $\lambda z := \rho\lambda\rho^{-1}z$, Z is projective of degree e. Indeed, using Proposition 5.3.3(ii) for $z,z' \in Z$, we have

$$K(z,\lambda z') = K(z,\rho\lambda\rho^{-1}z') = K_e(\rho^{-1}z,\lambda\rho^{-1}z') = \lambda^e K_e(\rho^{-1}z,\rho^{-1}z') = \lambda^e K(z,z'). \qquad \square$$

Proposition 5.8.7. *Let Z be a coherent space. Define on Z an equivalence relation $\equiv$ by*

$$z \equiv z' \quad \Leftrightarrow \quad K(z,z'') = K(z',z'') \quad \textit{for all } z'' \in Z.$$

Then the set $[Z]$ of equivalence classes

$$[z] := \{z' \in Z | z' \equiv z\} \quad (z \in Z)$$

is a nondegenerate coherent space with the coherent product

$$K([z],[z']) := K(z,z') \quad \textit{for all } z,z' \in Z. \tag{5.30}$$

The corresponding quantum spaces $\mathbb{Q}(Z)$ and $\mathbb{Q}([Z])$ are canonically isomorphic. In particular, if Z is projective, then $[Z]$ is projective, with scalar multiplication $\lambda[z] := [\lambda z]$.

Proof. Let Z be a coherent space and $z,z',w,w' \in Z$ with $[z] = [w]$ and $[z'] = [w']$. Then

$$K([z],[z']) = K(z,z') = K(w,z') = K(w,w') = K([w],[w']).$$

Thus, $K : [Z] \times [Z] \to \mathbb{C}$ is well-defined. It is straightforward to check that $([Z], K)$ is a coherent space. Now let $z, w \in Z$ with $K([z],[z']) = K([w],[z'])$ for all $z' \in Z$. Hence,

$$K(z,z') = K([z],[z']) = K([w],[z']) = K(w,z')$$

for all $z' \in Z$, giving $[z] = [w]$. Thus, $[Z]$ is nondegenerate. Let $T : \mathbb{Q}(Z) \to \mathbb{Q}([Z])$ be given by $\psi \to T\psi$, where $T\psi := \sum c_\ell |[z_\ell]\rangle$ for $\psi = \sum c_\ell |z_\ell\rangle \in \mathbb{Q}(Z)$. If $\psi = \sum c_\ell |z_\ell\rangle = 0$, then (for all $w \in Z$)

$$\langle [w]|T\psi = \sum c_\ell \langle [w]|[z_\ell]\rangle = \sum c_\ell K([w],[z_\ell]) = \sum c_\ell K(w,z_\ell) = 0.$$

Thus, $T\psi = 0$. Hence, $T : \mathbb{Q}(Z) \to \mathbb{Q}([Z])$ is a well-defined linear operator. Also, for $\psi \in \mathbb{Q}(Z)$, we have

$$\|T\psi\|^2 = \sum_j \sum_k \overline{c_j} c_k K([z_j],[z_k]) = \sum_j \sum_k \overline{c_j} c_k K(z_j,z_k) = \|\psi\|^2,$$

which implies that T is an isometry, hence injective. It is straightforward to see that T is surjective as well. Hence, T is an isomorphism.

If Z is projective, then $[Z]$ is projective with the same degree, with scalar multiplication $\lambda[z] := [\lambda z]$. Indeed, if Z is projective of degree e, we have

$$K([z],\lambda[z']) = K([z],[\lambda z']) = K(z,\lambda z') = \lambda^e K(z,z') = \lambda^e K([z],[z'])$$

for all $z, z' \in Z$ and $\lambda \in \mathbb{C}^\times$. □

Corollary 5.8.8. *Let Z be a projective coherent space. The canonical scalar multiplication on the nondegeneration space $[Z]$ is associative.*

Theorem 5.8.9. *Let Z be a coherent space, and let $A : Z \to Z$ be a coherent map with adjoint A^*. Then the class map $[A] : [Z] \to [Z]$, defined by*

$$[A][z] := [Az] \quad \textit{for all } z \in Z,$$

is a well-defined and coherent map with the unique adjoint $[A]^* = [A^*]$.

Proof. Let $z, z' \in Z$ with $[z] = [z']$. Using coherence of A, we have

$$K(Az,z'') = K(z,A^*z'') = K(z',A^*z'') = K(Az',z'')$$

for all $z'' \in Z$. Thus, $[Az] = [Az']$, and hence $[A] : [Z] \to [Z]$ is well-defined. Then, using coherence of A and applying the definition of the class map for the coherent maps A and A^*, we get

$$\begin{aligned} K([A][z],[z'']) &= K([Az],[z'']) = K(Az,z'') = K(z,A^*z'') \\ &= K([z],[A^*z'']) = K([z],[A^*][z'']) \end{aligned}$$

for all $z, z'' \in Z$. This guarantees that the class map $[A]$ is a coherent map with the unique adjoint $[A]^* = [A^*]$. □

Theorem 5.8.10. *Let Z be a coherent space. Then $[PZ] \cong P[Z]$, using a canonical identification. In particular,*
(i) *if Z is projective, then we have $[PZ] \cong [Z]$.*
(ii) *if Z is nondegenerate, then PZ is nondegenerate.*

Proof. The canonical map $\rho : [PZ] \to P[Z]$, given by $[(\lambda, z)] \to (\lambda, [z])$, is well-defined. It is also straightforward to check that ρ is a bijection. Let $[(\lambda, z)], [(\lambda', z')] \in [PZ]$. Then, we have

$$\begin{aligned} K_{\mathrm{pe}}(\rho[(\lambda,z)], \rho[(\lambda',z')]) &= K_{\mathrm{pe}}((\lambda,[z]),(\lambda',[z'])) \\ &= \overline{\lambda} K([z],[z'])\lambda' \\ &= \overline{\lambda} K(z,z')\lambda' = K_{\mathrm{pe}}((\lambda,z),(\lambda',z')) \\ &= K_{\mathrm{pe}}([(\lambda,z)],[(\lambda',z')]), \end{aligned}$$

implying that $\rho : [PZ] \to P[Z]$ is an isomorphism of coherent spaces. If Z is projective, then $PZ \cong Z$. Thus, we get $[PZ] \cong [Z]$. If Z is nondegenerate, then $[Z] \cong Z$. Hence, we have $[PZ] \cong P[Z] \cong PZ$, which implies that PZ is nondegenerate as well. □

5.9 Symmetries

This section introduces the concept of symmetries of coherent spaces, transformations of the space that preserve the coherent structure. More specifically, a symmetry of a coherent space Z is a bijection A of Z with the property that $K(z, Az') = K(A^T z, z')$ for another bijection A^T. More generally, a map $A : Z \to Z$ is called **coherent** if there is an **adjoint map** $A^* : Z \to Z$ such that

$$K(z, Az') = K(A^* z, z') \quad \text{for } z, z' \in Z. \tag{5.31}$$

If Z is nondegenerate, then the adjoint is unique, but not in general. A **symmetry** of Z is an invertible coherent map on Z with an invertible adjoint. Coherent maps form a semigroup $\mathrm{Coh}\, Z$ with identity; the symmetries form a subgroup. An **isometry** is a coherent map A that has an adjoint satisfying $A^* A = 1$. An invertible isometry is called **unitary**.

In the case of the trivial coherent product $K(z, z') := z^* z'$, equation (5.31) holds for every $n \times n$ matrix with conjugate transposition as the matrix adjoint. This motivates the general case, and shows in particular that the trivial coherent space has the **general linear group** $GL(n, \mathbb{C})$ of invertible complex $n \times n$ matrices as its group of symmetries. More generally, all invertible transformations of a Euclidean space $\mathbb{H}$ are symmetries of $Z = \mathbb{H}$, considered as a coherent space.

Example 5.9.1 (Distance regular graphs). The orbits of groups of linear self-mappings of a Euclidean space define coherent spaces with predefined transitive symmetry

groups. For example, the symmetric group Sym(5) acts as a group of Euclidean isometries on the 12 points of the icosahedron in $\mathbb{R}^3$. The coherent space, consisting of these 12 points with the induced coherent product, therefore has Sym(5) as a group of unitary symmetries. The quantum space is $\mathbb{C}^3$. The skeleton of the icosahedron is a distance-regular graph, here a double cover of the complete graph on six vertices. Many more interesting examples of finite coherent spaces are related to Euclidean representations of distance regular graphs and other highly symmetric combinatorial objects. See, for example, Brouwer et al. [55].

Example 5.9.2. The **Möbius space** $Z = \{z \in \mathbb{C}^2 \mid |z_1| > |z_2|\}$ is a coherent manifold with coherent product $K(z, z') := (\bar{z}_1 z_1' - \bar{z}_2 z_2')^{-1}$. A quantum space is the Hardy space of analytic functions on the complex upper half-plane with Lebesgue-integrable limit on the real line. The Möbius space has a large semigroup of coherent maps (a semigroup of compressions, Olshanski [225]) consisting of the matrices $A \in \mathbb{C}^{2\times 2}$ such that

$$\alpha := |A_{11}|^2 - |A_{21}|^2, \quad \beta := \overline{A}_{11}A_{12} - \overline{A}_{21}A_{22}, \quad \gamma := |A_{22}|^2 - |A_{12}|^2$$

satisfy the inequalities

$$\alpha > 0, \quad |\beta| \le \alpha, \quad \gamma \le \alpha - 2|\beta|.$$

It contains as a group of symmetries the group $GU(1,1)$ of matrices preserving the Hermitian form $|z_1|^2 - |z_2|^2$ up to a positive factor.

The compact symmetric spaces (and many noncompact ones) appear naturally as coherent spaces when equipped with a coherent product derived from the coherent states on semisimple Lie groups (see Perelomov [232]). In these cases, the coherent product is naturally related to the differential, metric, symplectic, and Kähler structure of the associated symmetric spaces (see Section 4.10 and Zhang et al. [316, Sections IIIC1 and VI]).

Highest weight representations. The example of the Möbius space generalizes to a large class of exactly solvable classical systems with finitely many degrees of freedom, corresponding to the coherent states from group representations discussed in Zhang et al. [316] and Simon [274], which are close to being computable (though not all needed details are in these papers). The constructions relate to central extensions of all **semisimple Lie groups** and associated **symmetric spaces** or **symmetric cones** and their **line bundles**. These provide many interesting examples of coherent manifolds. This follows from work on coherent states constructed from highest weight representations, discussed in monographs by Perelomov [232], by Faraut & Korányi [85], by Neeb [188].

Coherent states from highest weight representations induce on the corresponding coadjoint orbit a measure, a metric, a symplectic form, and an associated symplectic Poisson bracket. (See the Zhang et al. [316] survey for details from a physical point of

view.) The Poisson bracket defines a Lie algebra on phase space functions (C^∞ functions on the coadjoint orbit), hence an associated group of Hamiltonian diffeomorphisms, and the coherent state approach effectively quantizes this group. All this can be reconstructed directly from the associated coherent spaces. In particular, the nonclassical states of light in quantum optics called **squeezed states** are described by the Klauder spaces of Example 5.2.2, coherent spaces corresponding to the **metaplectic group**; see related work by Neretin [189].

5.10 Uses of coherent spaces

In this book, coherent spaces are used to define a coherent quantum physics. However, coherent spaces have many other uses. In general, they provide a setting for the study of geometry in a different direction than traditional metric, topological, and differential geometry. Just as it pays to study the properties of manifolds independently of their embedding into a Euclidean space, so does it appear fruitful to study the properties of coherent spaces independent of their embedding into a Hilbert space.

Coherent states are most often discussed as being parameterized by points on a connected manifold. But the concept of a coherent space also makes sense in a nontrivial way for finite spaces. There are strong relations between finite coherent spaces, finite metric spaces, graphs, and combinatorial designs. See Bekka & de la Harpe [29], Brouwer et al. [55], Godsil [101], Neumaier [190, 191, 192, 193]. This shows that the concept of coherent spaces provides a nontrivial extension of the theory of coherent states, in this respect similar to that of the measure-free coherent states of Horzela & Szafraniec [140].

Of particular importance is the use of reproducing kernels in complex analysis (see, for example, Faraut & Korányi, [85], Upmeier [287], and de Branges [50]) and group theory (see, for example, Neeb [188]), where they are the basis of many important theorems.

Reproducing kernel Hilbert spaces and the associated coherent states also have applications in many other fields of mathematics (see, for example, [2, 4, 5, 8, 9, 10, 34, 60, 161, 228, 232, 255, 256]), statistics and stochastic processes (see, for example, [37, 128, 129, 222, 228, 244, 256]), physics (see [2, 9, 10, 60, 93, 144, 161, 232]), and engineering (see [5, 93]).

In particular, there are relations to

(i) Christoffel–Darboux kernels for orthogonal polynomials;
(ii) Euclidean representations of finite geometries;
(iii) zonal spherical functions on symmetric spaces;
(iv) coherent states for Lie groups acting on homogeneous spaces;
(v) unitary representations of groups;
(vi) abstract harmonic analysis;
(vii) states of C^*-algebras in functional analysis;

(viii) reproducing kernel Hilbert spaces in complex analysis;
(ix) Pick–Nevanlinna interpolation theory;
(x) transfer functions in control theory;
(xi) positive definite kernels for radial basis functions;
(xii) positive definite kernels in data mining;
(xiii) positive definite functions in probability theory;
(xvi) exponential families in probability theory and statistics;
(xv) the theory of random matrices;
(xvi) Hida distributions for white noise analysis;
(xvii) Kähler manifolds and geometric quantization;
(xviii) coherent states in quantum mechanics;
(xix) squeezed states in quantum optics;
(xx) inverse scattering in quantum mechanics;
(xxi) Hartree–Fock equations in quantum chemistry;
(xxii) mean field calculations in statistical mechanics;
(xxiii) path integrals in quantum mechanics;
(xxiv) functional integrals in quantum field theory;
(xxv) integrable quantum systems.

These relations will be established elsewhere. The web site [200] will display at any time the most recent state of affairs.

6 Coherent quantum physics

Coherent quantum physics is quantum physics in terms of a coherent space and an appropriate coherent product. The kinematical structure and the meaning of the quantities are given by the symmetries (invertible coherent maps) of the coherent space.

If Z is a coherent manifold, a canonical symplectic form is often canonically determined by the coherent product through the coherent action principle discussed in Section 6.1. In a quantum mechanical context, Z is indeed a classical phase space or an extended phase space—typically a symplectic manifold, a Poisson manifold, or a circle or line bundle over such a manifold that incorporates the classical action variable (encoding the Berry phase under quantization). For example, the Aharonov–Bohm effect [3] needs the bundle formulation. A coherent space and its quantum space provide a classical and a quantum view of the same physical system, discussed in Section 6.2.

Section 6.3 shows that symmetries of coherent spaces can be quantized, that is, promoted to linear operators on the quantum space. This leads to quantum dynamics, which in special (completely integrable) situations can be solved in closed form in terms of classical motions on the underlying coherent space, if the latter has a compatible manifold structure. Close relations to concepts from geometric quantization and Kähler manifolds are pointed out in Section 6.4.

In order that the quantum space of a coherent space Z can describe physics, one needs not only a distinguished class of coherent states but a more general concept of state. In Sections 6.5–6.6, we give a general abstract setting for states in Lie $*$-algebras, emphasize the essential mathematical features and the close analogy between classical and quantum physics. We illustrate these generalities with a number of examples in Section 6.7. The coherent action principle is put to use in Section 6.8 for numerical quantum physics and an application to quantum chaos. Spectral issues can, in favorable cases, be handled in terms of dynamical Lie algebras, discussed in Section 6.9.

6.1 The coherent action principle

A basic approximation principle in quantum physics is the Dirac–Frenkel approach for reducing a nonrelativistic quantum problem to an associated approximate classical problem. In this approach, the variational principle for classical Lagrangian systems is rewritten for the present situation and then called the **Dirac–Frenkel variational principle**. It was first used by Dirac [69] and Frenkel [91], and found numerous applications; a geometric treatment is given in Kramer & Saraceno [166]. The action takes the form

$$I(\psi) = \int dt \psi^* (i\hbar\partial_t - H)\psi = \int dt (i\hbar\psi^*\dot{\psi} - \psi^* H\psi), \tag{6.1}$$

https://doi.org/10.1515/9783110667387-006

where the **quantum Hamiltonian** $H \in \mathrm{Lin}^{\times}\mathbb{H}$ is a self-adjoint operator. The coherent 1-form θ may be interpreted as the Lagrangian 1-form corresponding to the Dirac–Frenkel action. The Legendre transform of the **Lagrangian**

$$L(\psi) := i\hbar\psi^*\dot{\psi} - \psi^* H\psi$$

is the corresponding classical Hamiltonian

$$\langle H\rangle = \psi^* H\psi.$$

The Dirac–Frenkel action is stationary iff ψ satisfies the **Schrödinger equation**

$$i\hbar\dot{\psi} = H\psi.$$

If one has a coherent space Z and $\mathbb{H} = \mathbb{Q}(Z)$, a quantum space of Z, one can restrict ψ to coherent states, and we get the **coherent action**

$$I(z) = \int dt\langle z|(i\hbar\partial_t - H)|z\rangle$$

for the path $z(t)$. This **coherent action principle** was first proposed by Klauder [158]. The variational principle for the action $I(z)$ defines an approximate classical Lagrangian (and hence conservative) dynamics for the parameter vector $z(t)$. This **coherent dynamics** on Z is regarded as a semiclassical (or **semiquantal**) approximation of the quantum dynamics. In two important cases, the norm of the state is preserved by the coherent dynamics—Z must either be **normalized**, that is, $K(z,z) = 1$ for all $z \in Z$, or projective (as defined in Section 6.4). The approximation turns out to be exact when the Hamiltonian belongs to the infinitesimal Lie algebra of the symmetry group of the coherent state. It is inexact but good if it is not too far from such an element.

The classical problem created by the coherent action principle is again conservative, based an a classical action, which may or may not be transformable into an equivalent Hamiltonian problem. The latter depends on whether the Dirac–Frenkel Lagrangian is regular or singular. Thus, it is important that one understands the structure of classical singular Lagrangian problems. The regular case is characterized by the fact that the derivative of the Lagrangian 1-form, associated with the action, is nondegenerate, and hence defines a symplectic structure on the coherent space. For details, we refer to Kramer & Saraceno [166].

6.2 Systems with classical and quantum view

The symplectic structure on a coherent manifold, if present, provides a classical view of the system. On the other hand, as discussed in Section 5.3, the Moore–Aronszajn theorem implies that every coherent space Z has a quantum space $\mathbb{Q}(Z)$, unique up to

isometry, spanned algebraically by a distinguished set of coherent states $|z\rangle$ ($z \in Z$), satisfying

$$\langle z|z'\rangle = K(z,z') \quad \text{for } z,z' \in Z. \tag{6.2}$$

The antidual $\mathbb{Q}^\times(Z) := \mathbb{Q}(Z)^\times$ contains the Hilbert space completion $\overline{\mathbb{Q}}(Z)$ of $\mathbb{Q}(Z)$. In quantum mechanical applications, $\overline{\mathbb{Q}}(Z)$ is the Hilbert space containing the pure states, whereas $\mathbb{Q}^\times(Z)$ also contains unnormalizable wave functions. This provides a quantum view of the system.

Thus, coherent spaces allow both a classical and a quantum view of the same system. The two views are closely related, as the phase space points $z \in Z$ label the family of coherent states $|z\rangle$. Therefore, in some sense, the classical phase space and the quantum Hilbert space coexist in the framework of coherent spaces. The classical phase space is a quotient space of Z under the equivalence relation that identifies points, whose corresponding coherent states differ only by a scale factor. Thus, points in the phase space are in 1-1 correspondence with equivalence classes of points of Z, hence equivalence classes of labels of coherent states. The quantum space is the completion of the space spanned by all coherent states. It is a Hilbert space that can be realized as a space of functions on Z; the coherent states $|z\rangle$ are essentially the functions that map $z' \in Z$ to the coherent product $K(z,z')$.

The same abstract quantum system may allow different classical views. This is accommodated by writing the same Hilbert space in different but isomorphic ways as the quantum space of different coherent spaces.

The most conspicuous expression of this ambiguity is the **particle-wave duality**, a notion describing the seemingly paradoxical situation that the same quantum system may be approximately interpreted either in terms of classical particles or in terms of classical waves, though—depending on the circumstances—only one of the approximate views may be accurate enough to be useful.

If we regard a coherent space Z as a classical phase space, as often adequate, the functions

$$\widehat{\psi}(z) := \psi^* |z\rangle, \quad \psi \in \mathbb{Q}(Z)$$

are those classical phase space functions that have an immediate quantum meaning. Note that $\mathbb{Q}^\times(Z)$ consists of all complex-valued maps on Z that are continuous in the natural weak topology induced by the coherent product.

Constructing Hilbert spaces from a coherent space and its coherent product is much more flexible, and hence more powerful, than the standard approach of constructing Hilbert spaces from a function space and a measure on it. Virtually every Hilbert space arising in quantum mechanical practice can be neatly constructed as the quantum space of an appropriate coherent space; the examples in Section 5.7 give initial evidence of this.

The Glauber coherent states from Example 5.7.1(vi) are a particular instantiation of this concept. A more trivial case to keep in mind is to label all vectors in the finite-

dimensional Hilbert space $\mathbb{C}^n$, so that $Z = \mathbb{C}^n$ and $\langle z|z'\rangle = K(\overline{z}, z')$ with

$$K(z, z') := z^* z' = \sum_k \overline{z}_k z'_k. \tag{6.3}$$

This extends to infinite dimensions (the usual case in most of quantum physics) by replacing the sum by an appropriate integral, and shows that the traditional way of looking at Hilbert spaces can be fully accommodated with such a coherent space. However, this choice is poor from the point of view of the classical-quantum correspondence. As we shall see, there are far better choices, leading to a much increased flexibility compared to the traditional approach of defining Hilbert spaces by giving the inner product as a sum or integral. More importantly, as one works most of the time in Z and very little explicitly in the quantum space, one can often use classical intuition in quantum situations, and the economy of classical computations is often preserved.

Finite linear combinations of coherent states form a dense subspace $\mathbb{Q}(Z)$ of the Hilbert space $\overline{\mathbb{Q}}(Z)$. This implies that all quantum mechanical calculations, usually done in an orthonormal basis, can also be done on the basis of coherent states, and often far more efficiently. Most conceptual issues can be discussed in coherent terms, too. This makes the closeness to a classical description very plain, and removes much of the mystery of quantum physics.

The simplest classical systems have a finite number N of states, corresponding to a phase space Z with N elements. Their dynamics is that of a hopping process, a **continuous time Markov chain** determined by consistently specifying transition rates for hopping from one state to another. More complex classical systems have phase spaces Z that are finite-dimensional manifolds when there are only finitely many degrees of freedom. In particular, this is the arena of **classical mechanics** of point particles, where Z is a symplectic manifold, or—more generally—a Poisson manifold. The deterministic dynamics is defined on Z by Hamilton's equations, equivalently on phase space functions by means of the Poisson bracket. Finally, in **classical field theory**, the phase space Z is an infinite-dimensional space of fields in 3-dimensional space, the deterministic dynamics on Z is described by partial differential equations. Often an equivalent dynamics on phase space functions (now functions on fields) is given in terms of an appropriate Poisson structure on Z.

A 2-level quantum system also models a **spinning electron** in the ground state of its center of mass frame. Here the appropriate classical phase space is not the counterintuitive two state (up-down) model, which depends on a distinguished direction and hence sacrifices the spherical symmetry of the electron, but a 2-sphere in $\mathbb{R}^3$, the phase space of a classical spinning top. To account for the nonintegral spin of the electron, we should, in fact, take as classical phase space a circle bundle over the 2-sphere, given by the unit sphere in $\mathbb{C}^2$. This is related to the so-called **Hopf fibration**, a nontrivial topological object. The discussion of the Hopf fibration in terms of quaternions can be interpreted in terms of Pauli matrices, giving the traditional approach to 2-level systems. In terms of coherent states, all these technicalities are hidden—one has the

quantum space without having to bother about the latter. This economy of coherent states becomes more pronounced in more complicated models, which is the most important reason why they are useful.

6.3 Quantization and the dynamics of q-observables

This section introduces quantization procedures associated with symmetries of a coherent space. They generalize canonical transformations of a symplectic manifold, which is the special case of classical mechanics of point particles.

This leads to quantum dynamics, which in special (completely integrable) situations can be solved in closed form in terms of classical motions on the underlying coherent space, if the latter has a compatible manifold structure.

The importance of symmetries and more general coherent maps stems from the fact proved in Neumaier & Ghaani Farashahi [212] that there is a quantization operator Γ that associates with every coherent map A, a linear operator $\Gamma(A)$ on the quantum space $\mathbb{Q}(Z)$.[1]

Theorem 6.3.1 (Quantization theorem). *Let Z be a coherent space and $\mathbb{Q}(Z)$ a quantum space of Z. Then for any coherent map A on Z, there is a unique linear map $\Gamma(A) : \mathbb{Q}(Z) \to \mathbb{Q}(Z)$ such that*

$$\Gamma(A)|z\rangle = |Az\rangle \quad \textit{for all } z \in Z. \tag{6.4}$$

We call $\Gamma(A)$ the **quantization** of A, and Γ the **quantization map**. The quantization map furnishes a representation of the semigroup of coherent maps on Z (and hence of the symmetry group) on the quantum space of Z. In particular, this gives a **unitary representation** of the group of unitary coherent maps on Z.

The quantization operator is important as it reduces many computations with coherent operators in the quantum space of Z to computations in the coherent space Z itself. By the quantization theorem, large semigroups of coherent maps produce large semigroups of coherent operators, which may make complex calculations much more tractable. Coherent spaces with many coherent maps are often associated with symmetric spaces in the sense of differential geometry.

This essentially means that symmetries are those invertible linear transformations of the quantum space that map coherent states into coherent states, but is expressed without reference to the quantum space. This has very important implications for practical computations, reducing computations in the quantum space to simple computations in the coherent space. In particular, this makes certain problems easily exactly solvable that are in the traditional position or momentum representations nearly

1 In the literature, when applied to the special case, where $\overline{\mathbb{Q}}(Z)$ is a Fock space, $\Gamma(A)$ is often called the **second quantization** of A.

intractable. For example, the calculation of q-expectations requires in the traditional setting the evaluation of an integral over configuration space. In the case of field theory, the configuration space is infinite-dimensional, and already a rigorous definition of such integrals is very difficult. Moreover, finding closed formulas for integrals in high or infinite dimensions is more an art than a science. In contrast, in the coherent space approach, many q-expectations of interest can be obtained by differentiation, which is a fully algorithmic process.

Symmetries of a coherent space often represent the dynamical symmetries (see, for example, BARUT & RACZKA [27]) of an associated exactly solvable classical system. For example, if Z is a line bundle over a symplectic phase space, the symmetries would be all linear symplectic maps and their central extensions. Usually, only some of these preserve the Hamiltonian, and hence are symmetries of the system with this Hamiltonian.

For virtually all quantum systems of interest, there is a large classical **dynamical symmetry group**, which describes the symmetries of the underlying coherent space. Typically, this symmetry group is a (possibly infinite-dimensional) Lie group, much larger than the symmetry group of the system itself—which is the subgroup commuting with the Hamiltonian (in the nonrelativistic case) or preserving the action (in the relativistic case).

We now assume that Z is a **coherent manifold**. This means that Z carries a C^∞-manifold structure with respect to which the coherent product is smooth (C^∞). The relevant **observables** of the classical system are the discrete symmetries and the infinitesimal generators of the 1-parameter groups of symmetries that are smooth on the coherent product. They are promoted to q-observables of the corresponding quantum system through the quantization map. For a symmetry A, the corresponding q-observable is $\Gamma(A)$. For an **infinitesimal symmetry** X, that is, an element of the Lie algebra of generators of 1-parameter groups of the symmetry group, the corresponding **quantum symmetry**, acting on the quantum space of Z, is the q-observable given by the strong limit

$$d\Gamma(X) := \lim_{s\downarrow 0} \frac{\Gamma(e^{isX}) - 1}{is}.$$

Note that

$$d\Gamma(X + Y) = d\Gamma(X) + d\Gamma(Y), \quad e^{d\Gamma(X)} = \Gamma(e^X).$$

The quantization theorem, Theorem 6.3.1, may be regarded as a generalized **Noether principle** that automatically promotes all symmetries of Z to dynamical symmetries of the corresponding quantum system.

Thus, a coherent space contains intrinsically all information needed to interpret the quantum system, including that about which operators may be treated as q-observables.

The dynamics of a physical system is traditionally given by a **Hamiltonian**, a symmetric and Hermitian expression H in the q-observables. If the coherent space is in fact a **coherent manifold**, the **classical dynamics** determined by the Hamiltonian is given by a Poisson bracket canonically associated with the coherent space through variation of the coherent action discussed in Section 6.1. To get a Poisson bracket, in the case where the Lagrangian 2-form is degenerate, requires special measures; see Section 18.1 in the book by Neumaier & Westra [214]. Classical mechanics on Poisson manifolds, the most general setting for Hamiltonian dynamics in closed classical systems, is discussed in detail in Marsden & Ratiu [181]. Less general is classical mechanics on symplectic manifolds, and even more restricted is classical mechanics on cotangent bundles. The latter includes classical mechanics on the phase space $\mathbb{R}^{6N}$ for systems of N particles in Cartesian coordinates.

In a classical Hamiltonian system, the dynamics of a phase space function f is given by $\dot{f} = H \angle f$, where $f \angle g = \{g, f\}$ in terms of the Poisson bracket. For an N-particle system with particle positions q_j and particle momenta p_j, specializing this to $f = q_j$ and $g = p_j$, gives the classical equations of motion. In a quantum system, one has the same in the Heisenberg picture, and the resulting dynamics is equivalent to the Schrödinger equation in the Schrödinger picture.

Exactly solvable systems. In the special case, where the classical Hamiltonian is an infinitesimal symmetry of Z, and hence the quantum Hamiltonian has the form $\Gamma(H)$, the quantization lifts the classical phase-space trajectory to a quantum trajectory. Thus, if the Lie algebra of q-observables contains the Hamiltonian (and in some slightly more general situations), the quantum dynamics has the special feature that coherence is dynamically preserved. In terms of the Hamiltonian, the dynamics for pure quantum states ψ is traditionally given by the **time-dependent Schrödinger equation**

$$i\hbar \frac{d\psi}{dt} = d\Gamma(H)\psi$$

for the corresponding **quantum Hamiltonian** $d\Gamma(H)$. A dynamical symmetry preserved by H (in the classical case) or $d\Gamma(H)$ (in the quantum case) is a true symmetry of the corresponding classical or quantum system. The Fourier transform $\widehat{\psi}(E)$ satisfies the **time-independent Schrödinger equation**

$$d\Gamma(H)\widehat{\psi} = E\widehat{\psi}. \tag{6.5}$$

The following result shows that the solution of Schrödinger equations with a sufficiently nice Hamiltonian can be reduced to solving differential equations on Z.

Theorem 6.3.2. *Let Z be a coherent space, and let $\mathbb{G}$ be a Lie group of coherent maps with associated Lie algebra $\mathbb{L}$. Let $H(t) \in \mathbb{L}$ be a Hamiltonian with possibly time-dependent coefficients. Then the solution of the initial value problem*

$$i\hbar \frac{\partial}{\partial t}\psi_t = d\Gamma(H(t))\psi_t, \quad \psi_0 = |z_0\rangle \tag{6.6}$$

with $z_0 \in Z$ has for all times $t \geq 0$ the form of a coherent state, $\psi_t = |z(t)\rangle$ with the trajectory $z(t) \in Z$, defined by the initial value problem

$$i\hbar\dot{z}(t) = H(t)z(t), \quad z(0) = z_0.$$

This means that if a system is at some time in a coherent state, it will be at all times in a coherent state.

This conservation of coherence has the consequence that the quantum system is **exactly solvable**. This means that the complete solution of the dynamics of the quantum system can be reduced to the solution of the corresponding classical system. Effectively, the partial differential equations of quantum mechanics in the quantum space are solved in terms of ordinary differential equations on the underlying coherent space. In many cases, this implies that the spectrum can be determined explicitly in terms of the representation theory of the corresponding Lie algebras.

More generally (see, for example, Iachello [142]), we have an exactly solvable system whenever the Hamiltonian $H(t)$ is a linear combination of infinitesimal symmetries with coefficients given by **Casimirs** of the Lie algebra $\mathbb{L}$ of infinitesimal symmetries. That is, in the classical case, central elements of the Lie–Poisson algebra $C^\infty(\mathbb{L}^*)$, and in the quantum case, the universal enveloping algebras of $\mathbb{L}$. On any orbit of the symmetry group, these Casimirs are represented by multiplication with a constant. One can therefore extend the coherent space Z without changing the quantum space by treating the corresponding multiples of the coherent states as new coherent states of an extended coherent space, whose elements are labeled by pairs of elements of Z and appropriate multipliers. This turns the algebra of Casimirs into an Abelian group of symmetries of the extended coherent space, which, together with original symmetries, provides an action of a central extension of the original symmetry group as a symmetry group of the extended coherent space.

6.4 Relations to geometric quantization

Often, classical symmetries are promoted to quantum symmetries in a projective representation. Then the symmetry group of the extended phase space is a proper central extension of the symmetry group of the original space. It acts on an extended phase space, whose dimension is larger. For example, the classical phase space with n spatial degrees of freedom has dimension $2n$, but the associated Heisenberg algebra, the central extension of an Abelian group with $2n$ generators, has dimension $2n+1$, as the canonical commutation relations for the extended Poisson bracket (or in the quantum case for the commutator) require an additional central generator.

Such a central extension is the rule rather than the exception. The extra dimension, often discussed in the context of the **Berry phase** or **geometric phase**, accounts for topological features, such as the Aharonov–Bohm effect. But it also occurs in classical physics. For example, a classical electromagnetic field exhibits

topological effects when the field strength is not globally integrable to a vector potential.

The explicit description of a central extension in terms of the original symmetry group involves so-called **cocycles**. Rather than with the original symmetry group, one can indeed work directly with a central extension of the group, acting on the extended phase space. (Examples where this works are the Möbius space and the Klauder spaces discussed before.) In this way, one can avoid the use of cocycles, as the relevant projective representations become ordinary representations of the central extension. Thus, the extended description generally reflects the quantum properties in a more symmetric way than the original coherent space.

In our present setting, the extended phase space is modeled by a projective coherent space; see Section 5.8. Projectivity is typically needed when one wants to have all symmetries of interest represented coherently. In particular, projective coherent spaces coherently represent central extensions of groups in cases where the original group is represented by a projective representation that would lead to coherent maps only up to additional scalar factors called cocycles.

Geometrically, the extended phase space takes the form of a line bundle. In the case of the Heisenberg algebra, the line bundle is trivial, formed by $Z = \mathbb{C} \times \mathbb{C}^n$ with componentwise conjugation, scalar multiplication defined by $\alpha(\lambda, s) := (\alpha\lambda, s)$, and coherent product

$$K(z, z') := \lambda\lambda' e^{s^T s'/\hbar} \quad \text{for } z = (\lambda, s),\ z' = (\lambda', s').$$

The foregoing give rise to a projective coherent space Z, whose quantum space $\mathbb{Q}(Z)$ is the bosonic Fock space with n independent oscillators, and the coherent states are the multiples of the Glauber coherent states. Indeed, the coherent states

$$|\lambda, s\rangle, \quad \lambda, s \in \mathbb{C}$$

in a single-mode Fock space have the Hermitian inner product

$$\langle \lambda, s | \lambda', s' \rangle = \bar{\lambda}\lambda' e^{\bar{s}s'/\hbar}.$$

The Klauder spaces from Example 5.2.2 extend this construction to an arbitrary finite or infinite number of modes. In terms of the traditional Fock space description, the coherent states are the simultaneous eigenstates of the annihilator operators,

$$a|z\rangle = \mathbf{z}|z\rangle \quad \text{for } z \in Z.$$

More generally, (see Neumaier & Ghaani Farashahi [212]), Klauder spaces provide an elegant and efficient approach to the properties of creation and annihilation operators.

A coherent space generalizes finite-dimensional symplectic manifolds with a polarization that induces a complex structure on the manifold. A projective coherent

space generalizes a corresponding **Hermitian line bundle** Z, that is, a line bundle with a Hermitian connection. Such line bundles are usually discussed in the context of geometric quantization. **Geometric quantization** (see, for example, Engliš [82, 83], Schlichenmaier [263], Bar-Moshe & Marinov [25])proceeds from a symplectic manifold $\mathbb{K}$. It constructs (in the group case in terms of integral cohomology) a polarization that defines a Hermitian line bundle $Z = \mathbb{C}\mathbb{K}$ and an associated **Kähler potential** (which is essentially the logarithm of the coherent product). This potential turns $\mathbb{K}$ into a Kähler manifold with a natural Kähler metric, Kähler measure, and symplectic Kähler bracket. If the Kähler metric is definite (which is always the case if Z is a compact symmetric space), there is an associated Hilbert space of square integrable functions, on which quantized operators can be defined by a recipe of van Hove [141].

An **involutive coherent manifold** is a coherent manifold Z equipped with a smooth mapping that assigns to every $z \in Z$ a **conjugate** $\bar{z} \in Z$ such that $\overline{\bar{z}} = z$ and $\overline{K(z,z')} = K(\bar{z},\bar{z}')$ for $z, z' \in Z$. Under additional conditions, an involutive coherent manifold carries a canonical Kähler structure, turning it into a **Kähler manifold**. For semisimple finite-dimensional Lie algebras, the irreducible highest weight representations have nice coherent space formulations. In the literature, the logarithm of the coherent product figures under the name of Kähler potential. Zhang et al. [316] relate the latter to coherent states. The coherent quantization of Kähler manifolds is equivalent to traditional **geometric quantization of Kähler manifolds**. But in the coherent setting, quantization is not restricted to finite-dimensional manifolds, which is important for quantum field theory.

The coherent product and the conjugation are C^∞-maps on the line bundle. For this line bundle to exist, the symplectic manifold must also carry a positive definite Kähler potential $F : Z \times Z \to \mathbb{C}$, satisfying a generalized **Bohr–Sommerfeld quantization condition** defined by the integrality of some cohomological expression. In this case, Z is a projective coherent space with coherent product

$$K(z,z') := e^{-F(z,z')}.$$

The quantum space of the projective coherent space carries the representation satisfying the conditions of a successful geometric quantization. This procedure, called **Berezin quantization**, is the most useful way of performing geometric quantization; see, for example, Schlichenmaier [263].

6.5 Lie $*$-algebras

In order that the quantum space of a coherent space Z can describe physics, one needs not only a distinguished class of coherent states but a more general concept of state. In the remainder of this chapter, we give a generalized abstract setting for states in both

classical and quantum mechanics to emphasize the essential mathematical features and the close analogy between classical and quantum physics.

A (complex) **Lie algebra** is a complex vector space $\mathbb{L}$ with a distinguished **Lie product**, a bilinear operation on $\mathbb{L}$, satisfying $X \between X = 0$ for $X \in \mathbb{L}$, and the **Jacobi identity**

$$X \between (Y \between Z) + Y \between (Z \between X) + Z \between (X \between Y) = 0 \quad \text{for } X, Y, Z \in \mathbb{L}.$$

A **Lie $*$-algebra** is a complex Lie algebra $\mathbb{L}$ with a distinguished element $1 \neq 0$, called **one**, and a mapping $*$ that assigns to every $X \in \mathbb{L}$ an **adjoint** $X^* \in \mathbb{L}$ such that

$$\begin{aligned} &(X + Y)^* = X^* + Y^*, \quad (X \between Y)^* = X^* \between Y^*, \\ &X^{**} = X, \quad (\lambda X)^* = \lambda^* X^*, \\ &1^* = 1, \quad X \between 1 = 0 \end{aligned}$$

for all $X, Y \in \mathbb{L}$ and $\lambda \in \mathbb{C}$ with complex conjugate λ^*. We identify the multiples of 1 with the corresponding complex numbers.

A **state** on a Lie $*$-algebra $\mathbb{L}$ is a positive semidefinite Hermitian form $\langle \cdot, \cdot \rangle$, antilinear in the first argument and normalized such that $\langle 1, 1 \rangle = 1$.

A group G **acts** on a Lie $*$-algebra $\mathbb{L}$ if for every $A \in G$ there is a linear mapping that maps $X \in \mathbb{L}$ to $X^A \in \mathbb{L}$ such that

$$\begin{aligned} &(X \between Y)^A = X^A \between Y^A, \\ &\left(X^A\right)^B = X^{AB}, \quad \left(X^A\right)^* = \left(X^*\right)^A, \quad X^1 = X, \quad 1^A = 1 \end{aligned}$$

for all $X, Y \in \mathbb{L}$ and all $A, B \in G$. Thus, the mappings $X \to X^A$ are $*$-automorphisms of the Lie $*$-algebra. Such a family of mappings is called a **unitary representation** of G on $\mathbb{L}$.

Often, unitary representations arise by writing the Lie $*$-algebra $\mathbb{L}$ as a vector space of complex $n \times n$ matrices, closed under conjugate transposition * and commutation with $X \between Y := \frac{i}{\hbar}[X, Y]$, and G as a group of unitary $n \times n$ matrices such that $X^A := A^{-1} X A \in \mathbb{L}$ for all $X \in \mathbb{L}$.

6.6 Quantities, states, uncertainty

In classical and quantum physics, **physical systems** are modeled by appropriate Lie $*$-algebras $\mathbb{L}$, whose elements are interpreted as the **quantities** of the system modeled. Each physical system may exist in different **instances**; each instance specifies a particular system under particular conditions. A state defines the **properties** of an instance of a physical system described by a model, and hence what **exists** in the system. Properties depend on the state and are expressed in terms of definite but uncertain values of the quantities:

(GUP) General uncertainty principle: *In a given state, any quantity $X \in \mathbb{L}$ has the* ***uncertain value***

$$\overline{X} := \langle X \rangle := \langle 1, X \rangle \tag{6.7}$$

with an ***uncertainty*** *of*[2]

$$\sigma_X := \sqrt{\langle X - \overline{X}, X - \overline{X} \rangle} = \sqrt{\langle X, X \rangle - |\overline{X}|^2}. \tag{6.8}$$

Through (6.7), each state induces an element $\langle \cdot \rangle$ of the **dual** of $\mathbb{L}$, the space $\mathbb{L}^*$ of linear functionals on $\mathbb{L}$.

The identification of formal properties given by uncertain values with real life properties of a physical system, is done (see the discussion in Section 8.3) by means of

(CC) Callen's criterion (Callen [58, p. 15]): *Operationally, a system is in a given state if its properties are consistently described by the theory for this state.*

This is enough to find out in each single case how to approximately measure the uncertain value of a quantity of interest, though it may require considerable experimental ingenuity to do so with low uncertainty. The uncertain value $\overline{X}$ is considered informative only when its uncertainty σ_X is much less than $|\overline{X}|$.

As position coordinates are dependent on a convention about the coordinate system used, so all system properties are dependent on the **conventions** under which they are viewed. To be objective, these conventions must be interconvertible. This is modeled by a group G of **symmetries** acting transitively both on the spacetime manifold M considered and on the set W of conventions. We write these actions on the left, so that $A \in G$ maps $x \in M$ to Ax, and $w \in W$ to Aw.

To be applicable to a physical system, a representation of G on the Lie $*$-algebra $\mathbb{L}$ of quantities must be specified. Depending on the model, this representation accounts for conservative dynamics and the principle of relativity in its nonrelativistic, special relativistic, or general relativistic situation. It also caters for the presence of internal symmetries of a physical system. Correspondingly, G may be a group of matrices, a Heisenberg group, the Galilei group, the Poincaré group, or a group of volume-preserving diffeomorphisms of a spacetime manifold M.

A **particular physical system** in all its views is described by a family of states $\langle \cdot, \cdot \rangle_w$ indexed by a convention $w \in W$, satisfying the covariance condition

$$\langle X, Y \rangle_{Aw} = \langle X^A, Y^A \rangle_w \tag{6.9}$$

2 Since the state is positive semidefinite, the first expression shows that σ_X is a nonnegative real number. The equivalence of both expressions defining σ_X follows from $\langle X \rangle = \overline{X}$ and

$$\begin{aligned} \langle X - \overline{X}, X - \overline{X} \rangle &= \langle X, X \rangle - \langle X, \overline{X} \rangle - \langle \overline{X}, X \rangle + \langle \overline{X}, \overline{X} \rangle \\ &= \langle X, X \rangle - \langle X \rangle^* \overline{X} - \overline{X}^* \langle X \rangle + |\overline{X}|^2 = \langle X, X \rangle - |\overline{X}|^2. \end{aligned}$$

for $X, Y \in \mathbb{L}$, $w \in W$, $A \in G$. In particular, uncertain values transform as

$$\langle X \rangle_{Aw} = \langle X^A \rangle_w. \tag{6.10}$$

A **subsystem** of a particular physical system is defined by specifying a Lie $*$-subalgebra and restricting the family of states to this subalgebra.

If (as is commonly done) we work within a fixed affine coordinate system in a spacetime (homeomorphic to some) $\mathbb{R}^d$, the only conditions relevant are when and where a system is described; all other conditions are handled implicitly by covariance considerations. In this case, W is simply the spacetime M, and G is the group of affine translations $T_z : x \to x + z$ of M by z. In this case,

$$\overline{X}(x) = \langle X \rangle_x, \quad \sigma_X(x) = \sqrt{\langle X, X \rangle_x - \left|\overline{X}(x)\right|^2}$$

define the **value** $\overline{X}(x)$ of X at x and its **uncertainty** $\sigma_X(x)$ at x, and (6.9) and (6.10) become

$$\langle X, Y \rangle_{x+z} = \langle X^{T_z}, Y^{T_z} \rangle_x, \quad \langle X \rangle_{x+z} = \langle X^{T_z} \rangle_x.$$

The value $\overline{X}(x)$ is (in principle) **observable** with **resolution** $\delta > 0$ if it varies slowly with x and has a sufficiently small uncertainty. More precisely, if Δ denotes the set of spacetime shifts that are imperceptible in the measurement context of interest, observability with resolution δ requires that

$$\begin{aligned} &\left|\overline{A}(x + h) - \overline{A}(x)\right| \leq \delta \quad \text{for } h \in \Delta, \\ &\sigma_X(x) \ll \left|\overline{A}(x)\right| + \delta. \end{aligned}$$

We require that the translation group is generated by a covariant **momentum vector** $p \in \mathbb{L}^d$ with Hermitian components, in the sense that

$$\frac{\partial}{\partial x_\nu} X^{T_x} = p_\nu \angle X \tag{6.11}$$

for $X \in \mathbb{L}$, $x \in M$, and all indices ν. From the covariance condition (6.9), we conclude that

$$\frac{\partial}{\partial x_\nu} \langle X, Y \rangle_x = \langle p_\nu \angle X, Y \rangle_x + \langle X, p_\nu \angle Y \rangle_x. \tag{6.12}$$

In particular, the uncertain values satisfy the **covariant Ehrenfest equation**

$$\frac{\partial}{\partial x_\nu} \langle X \rangle_x = \langle p_\nu \angle X \rangle_x. \tag{6.13}$$

In classical or quantum multiparticle mechanics (as opposed to field theory), space and time are treated quite differently, and we are essentially in the case $d = 1$ of the

above, where the convention about views of system properties is completely specified by the time $t \in \mathbb{R}$. In this case, the above specializes to

$$\overline{X}(t) = \langle X \rangle_t, \quad \sigma_X(t) = \sqrt{\langle X, X \rangle_t - \left|\overline{X}(t)\right|^2}.$$

The time translation group is generated by a Hermitian **Hamiltonian** $H \in \mathbb{L}$, so that

$$\frac{d}{dt} X^{T_t} = H \angle X. \tag{6.14}$$

$$\frac{d}{dt} \langle X, Y \rangle_t = \langle H \angle X, Y \rangle_t + \langle X, H \angle Y \rangle_t. \tag{6.15}$$

In particular, the uncertain values satisfy the **Ehrenfest equation**

$$\frac{d}{dt} \langle X \rangle_t = \langle H \angle X \rangle_t, \tag{6.16}$$

providing a deterministic dynamics for the q-expectations. This generalizes the Ehrenfest equation discussed in Section 2.2.

6.7 Examples

1. A simple classical example is $\mathbb{L} = \mathbb{C}^3$, with the cross product as Lie product. It is isomorphic to the Lie algebra $so(3, \mathbb{C})$, and describes in this representation a **rigid rotator**. The dual space $\mathbb{L}^*$ is spanned by the three components of J, and the functions of J^2 are the Casimir operators. Assigning to J a particular 3-dimensional vector with real components (since J has Hermitian components) gives the classical angular momentum in a particular state.
2. The same Lie algebra is also isomorphic to $su(2)$, the Lie algebra of traceless Hermitian 2×2 matrices, and then describes a single **qubit**. In this case, we think of $\mathbb{L}^*$ as mapping the three Hermitian Pauli matrices σ_j to three real numbers S_j, and extending the map linearly to the whole Lie algebra. Augmented by $S_0 = 1$ to account for the identity matrix, which extends the Lie algebra to that of all Hermitian matrices, this leads to the classical description of the qubit discussed in Section 8.6.
3. Consider the Lie $*$-algebra $\mathbb{L}$ of smooth functions $f(p, q)$ on classical phase space with the negative Poisson bracket as Lie product and $*$ as complex conjugation. Given a $*$-homomorphism ω with respect to the associative pointwise multiplication, determined by the classical values $\overline{p}_k := \omega(p_k)$ and $\overline{q}_k := \omega(q_k)$, the states defined by

$$\langle X, Y \rangle := \omega(X^* Y)$$

reproduce classical deterministic dynamics. More generally, $\mathbb{L}$ can be partially ordered by defining $f \geq 0$ iff f takes values in the nonnegative reals. Given a monotone $*$-linear functional ω on $\mathbb{L}$, satisfying $\omega(X^*) = \omega(X)^*$ for $X \in \mathbb{L}$ and $\omega(1) = 1$, the states defined by

$$\langle X, Y\rangle := \omega(X^* Y)$$

reproduce **classical stochastic dynamics** in the Koopman picture discussed in Section 7.10. In both cases, $\langle X\rangle = \omega(X)$.

4. The basic example of interest for **isolated quantum systems** is the Lie $*$-algebra $\mathbb{L}(Z)$ of linear operators acting on the quantum space $\mathbb{H} = \mathbb{Q}(Z)$ of the coherent space Z, with Lie product

$$X \angle B := \frac{i}{\hbar}[X, B] = \frac{i}{\hbar}(XB - BX). \tag{6.17}$$

Note that if no quantities are distinguished, Z s just a Euclidean space $\mathbb{H}$ and $\mathbb{L}(Z) = \operatorname{Lin}\mathbb{H}$.

The action of the translation group on $X \in \mathbb{L}$ is given by

$$X^{T_x} := U(x)^* X U(x)$$

with unitary operators $U(x)$ satisfying $U(0) = 1$ and $U(x)U(y) = U(x+y)$. The states of interest are the **regular states**, defined by

$$\langle X, Y\rangle_x = \operatorname{Tr}\big(Y\rho(x)X^*\big)$$

for some positive semidefinite Hermitian **density operator** $\rho(x) \in \overline{\mathbb{Q}}(Z)$ with $\operatorname{Tr}\rho(x) = 1$. In this case, the uncertain values

$$\langle X\rangle_x = \operatorname{Tr}\big(X\rho(x)\big)$$

viewed from $x \in \mathbb{R}^d$ are the **q-expectations**[3] of X, and the uncertainty can be expressed in terms of q-expectations, too. For Hermitian X, it is given by

$$\sigma_X(x) = \sqrt{\langle X^2\rangle_x - \langle X\rangle_x^2}.$$

6.8 Coherent chaos

The coherent action principle is the basis of much of traditional numerical quantum mechanics, which heavily relies on variational methods. It plays an important role in

3 Traditionally, $\langle X\rangle$ is called the expectation value of X, but such a statistical interpretation is not needed and is not even possible when X is defective, hence has no spectral resolution.

approximation schemes for the dynamics of quantum systems. In many cases, a viable approximation is obtained by restricting the state vectors $\psi(t)$ to a linear or nonlinear manifold of easily manageable states $|z\rangle$ parameterized by classical parameters z, which can often be given a physical meaning.

What is commonly called a mean field theory is the simplest coherent state approximation. This is already much better than a classical limit view, and—in particular—corrects for the missing zero point energy terms in the latter.

An important application of this situation are the **time-dependent Hartree–Fock equations** (see, for example, McLachlan & Ball [183]), obtained by choosing Z to be a Grassmann space.[4] This gives the Hartree–Fock approximation, which is at the heart of dynamical simulations in quantum chemistry. It can usually predict energy levels of molecules to within 5 % accuracy. Choosing Z to be a larger space (obtained by the methods of Section 5.6) enables one to achieve accuracies approaching 0.001 %.

Apart from Hartree-Fock calculations (symmetry group $U(N)$ on coadjoint orbits of Slater states), this covers Hartree–Fock–Bogoliubov methods (see, for example, Goodman [102]), which include Bogoliubov transformations to get a quasiparticle picture (symmetry group $SO(2N)$), and Gaussian methods (see, for example, Pattanayak & Schieve [230], Ono & Ando [226]) used in quantum chemistry (symmetry group $ISp(2N)$). There are time-dependent versions of these, and extensions that go beyond the mean field picture, using either Hill-Wheeler equations in the generator coordinate method (see, for example, Griffin & Wheeler [108]) or coupled cluster expansions (see, for example, Bartlett & Musial [26]) around the mean field.

We now show that the coherent action principle reveals how chaos emerges through coarse-graining from the exact quantum dynamics, in spite of the linearity of the Schrödinger equation.

Zhang & Feng [315] use the Dirac–Frenkel variational principle restricted to general coherent states to get a semiquantal system of ordinary differential equations approximating the dynamics of the q-expectations of macroscopic operators of certain multiparticle systems. At high resolution, this deterministic dynamics is highly chaotic. This chaoticity is a general feature of approximation schemes for the dynamics of q-expectations or the associated reduced density functions.

Zhang and Feng derive in a purely mathematical way—without referring to probability or statistics—the equations that they show to be chaotic. Thus, what they do is completely independent of any particular interpretation of quantum physics. They construct a semiclassical dynamics (where the relevant operators are replaced by their q-expectations) and then discuss the resulting system of ordinary differential equa-

4 A Grassmann space is a manifold of all k-dimensional subspaces of a vector space. It is one of the symmetric spaces.

tions. It turns out to be chaotic. The exact quantum dynamics would be given instead by partial differential equations!

In the overview of their paper, Zhang & Feng [315, pp. 4–9] state that they focus attention on understanding the question of quantum-classical correspondence (QCC), the search for an unambiguous classical limit, starting purely from quantum theory. They explore how, under suitable conditions, classical chaos can emerge naturally from quantum theory. They use the semiquantal method discussed above for the exploration of the correspondence between quantum and classical dynamics and quantum nonintegrability. They mention the relations to geometric quantization and coherent states, and work in a group theoretic setting corresponding to coherent states defined by coadjoint orbits of semisimple Lie groups. Their coset space G/H (or rather a complex line bundle over it arising in geometric quantization and carrying some of the phase information) discussed in [315, pp. 39] is a coherent space with coherent product given by their (3.1.8). The variation of the effective quantum action in their (3.2.11) is the coherent action principle. The result of the variation is (in the regular case) a symplectic system of differential equations that has a semiclassical (or, as they say, semiquantal) interpretation. This system gives an approximate dynamics for the q-expectations of the generators of the dynamical group. This dynamics is chaotic when the classical limit of the quantum system is not integrable.

6.9 Dynamical Lie algebras

The time-independent Schrödinger equation (6.5) generalizes easily to a more general **implicit Schrödinger equation**

$$I(E)\psi = 0 \tag{6.18}$$

with an energy-dependent **system operator** $I(E)$, and ψ in the antidual of some Euclidean space $\mathbb{H}$. This more general formulation fits naturally the coherent space setting, and everything said so far (corresponding to $I(E) = E - d\Gamma(H)$) generalizes to the general implicit formulation.

A nonlinear $I(E)$ typically appears in reduced effective descriptions of systems derived from a more complicated Hamiltonian setting and in relativistic systems. (The antidual is needed to account for a possible continuous spectrum.)

This section discusses implicit Schrödinger equations for the exactly solvable case, where the system operator $I(E)$ is contained in a Lie algebra $\mathbb{L}$ with known representation theory. This is the setting, where a tractable dynamical symmetry group for the Hamiltonian is known and covers many interesting systems.

For example, the system operator $I(E) = p_0^2 - \mathbf{p}^2 - (mc)^2$ with $p_0 = E/c$ describes a free spin 0 particle. This generalizes to a quadratic implicit Schrödinger equation

$$\left(\pi^2 - \frac{ige\hbar}{c}\mathbf{S}\cdot\mathbf{F}(x) - (mc)^2\right)\psi = 0 \tag{6.19}$$

for a particle of charge e, mass m, and arbitrary spin in an electromagnetic field. Here

$$\pi = \begin{pmatrix} \pi_0 \\ \boldsymbol{\pi} \end{pmatrix} := p + e\hbar A(x) \tag{6.20}$$

is a gauge-invariant 4-vector, $\mathbf{S}$ is the 3-dimensional spin vector representing the intrinsic angular momentum of a particle of spin $j = 0, \frac{1}{2}, 1, \ldots$, the 3-vector $\mathbf{F}(x) = \mathbf{E}(x) + ic\mathbf{B}(x)$ is the Riemann–Silberstein vector encoding the electric field $\mathbf{E}(x)$ and the magnetic field $\mathbf{B}(x)$, and g is the dimensionless **g-factor** of the **magnetic moment**

$$\boldsymbol{\mu}_s := -\frac{g\mu_B}{\hbar}\mathbf{S},$$

where μ_B is a constant called the **Bohr magneton**, and ψ is a wave function with $s = 2j + 1$ components. For spin $j = 1/2$, we have $s = 2$ components. Hence, being second order, 4 local degrees of freedom, corresponding to the 4 components of the (first-order) Dirac equation, which is equivalent to the special case $g = 2$.

In the special case, where the dependence on E is linear, we have

$$I(E) = EM - N \tag{6.21}$$

with fixed $M, N \in \mathbb{L}$. This covers the simple case of a harmonic oscillator, where $M = 1$, $N = \frac{1}{2}(p^2/m + Kq^2)$ is the Hamiltonian, and the Lie algebra is the **oscillator algebra**, with generators $1, p, q, H$ (or, in complex form, $1, a, a^*, a^*a$). It also covers a family of practically relevant exactly solvable systems with Lie algebra $\mathbb{L} = so(2,1) \oplus \mathbb{C} = su(1,1) \oplus \mathbb{C}$ discussed in detail in the book by Wybourne [310], containing among others the case of a particle of mass m in a Coulomb field, with $M = r = |\mathbf{q}|$ and $N = MH$, where

$$H = \frac{1}{2}m\mathbf{v}^2 - \frac{\alpha}{|\mathbf{q}|}$$

is the Coulomb Hamiltonian.

If $I(E)$ belongs for all E to some Lie $*$-algebra $\mathbb{L}$ acting in a (reducible or irreducible) representation, then $\mathbb{L}$ is called a **dynamical Lie algebra**[5] of the problem.

In general, the requirement for a dynamical symmetry group is just that all quantities of physical interest in the system can be expressed in the Lie–Poisson algebra (in the classical case) or the universal enveloping algebra (in the quantum case) of the corresponding Lie algebra. In this case, the label "dynamical" is a misnomer, and **kinematic symmetry group** would be more appropriate. The kinematic symmetry

5 One can always take the dynamical Lie algebra to be the Lie algebra $\mathrm{Lin}\mathbb{H}$ of all linear operators on the nuclear space $\mathbb{H}$. For this choice, the dynamical Lie algebra offers no advantage over the standard treatment. Therefore, it is usually understood that the dynamical Lie algebra is much smaller than $\mathrm{Lin}\mathbb{H}$, although mathematically there is no such restriction.

group is an integral part of the Hamiltonian or Lagrangian setting; so one usually gets it directly from the formulation and a look at the obvious symmetries. For any anharmonic oscillator, it is $Sp(2)$; for any system of N particles in $\mathbb{R}^3$, it is the symplectic group $Sp(6N)$, generated by the inhomogeneous quadratics in p and q.

If a problem has a dynamical symmetry group such that the (discrete or continuous) spectrum of all elements of its Lie algebra $\mathbb{L}$ is exactly computable, then the spectrum of the system can be found exactly. In the best understood cases, $\mathbb{L}$ is a finite-dimensional semisimple Lie algebra. Here everything is tractable more or less explicitly since the representation theory of these Lie algebras and their corresponding groups is fully understood. A problem solvable in this way is called **integrable**.

The **spectrum** of the nonlinear eigenvalue problem (6.18) is the set $\operatorname{Spec} I$ of all $E \in \mathbb{C}$ such that $I(E)$ is not invertible. In terms of (generalized) eigenvalues and eigenvectors of the nonlinear eigenvalue problem,

$$I(E)|\xi, E\rangle = \lambda(\xi, E)|\xi, E\rangle,$$

where ξ is a label distinguishing different eigenvectors $|\xi, E\rangle$ in a (generalized) orthonormal basis of the eigenspace corresponding to the eigenvalue E. To cover the continuous spectrum (where eigenvectors are unnormalized, hence do not belong to the Hilbert space), we work in a Euclidean space $\mathbb{H}$, on which the Hamiltonian acts as a linear operator. The Hilbert space of the problem is then the completion $\overline{\mathbb{H}}$ of this space, and $\mathbb{H} \subseteq \overline{\mathbb{H}} \subseteq \mathbb{H}^\times$ is a Gelfand triple. Therefore,

$$I(E)|\xi, E\rangle = 0 \quad \text{whenever } \lambda(\xi, E) = 0.$$

Thus,

$$\operatorname{Spec} I = \{E \in \mathbb{R} \mid \lambda(\xi, E) = 0 \text{ for some } \xi \in \operatorname{Spec} I(E)\}.$$

Moreover, it is easy to see that all eigenvectors of the nonlinear eigenvalue problem have the form $|\xi, E\rangle$. Therefore, the spectrum is given by the set of solutions of the nonlinear equation $\lambda(\xi, E) = 0$.

In many cases of interest (for example, see Section 6.5, when $\mathbb{L}$ is a Lie $*$-algebra), $\mathbb{L} = \mathbb{L}_0 \oplus \mathbb{C}$; then we may write

$$I(E) := m(E)X(E) - k(E), \tag{6.22}$$

where $m(E)$ and $k(E)$ are scalars not vanishing simultaneously, and $X(E) \in \mathbb{L}_0$. If

$$X(E)|\xi, E\rangle = \xi|\xi, E\rangle$$

is a complete system of (generalized) eigenvalues and eigenvectors of $X(E)$, then

$$I(E)|\xi, E\rangle = \lambda(\xi, E)|\xi, E\rangle, \quad \lambda(\xi, E) = m(E)\xi - k(E). \tag{6.23}$$

Therefore,

$$I(E)|\xi, E\rangle = 0 \quad \text{whenever } \lambda(\xi, E) = 0.$$

Again, all eigenvectors of the nonlinear eigenvalue problem have the form $|\xi, E\rangle$, and the spectrum is given by the set of solutions of $\lambda(\xi, E) = 0$.

In the integrable case, one may find the states $|\xi, E\rangle$ by transforming $I(E)$ to elements from a standard set of representatives of the conjugacy classes, and has to work out explicit spectral factorizations for these. For semisimple Lie algebras $\mathbb{L}$ in finite dimensions, each Lie algebra element is in a Cartan subalgebra, and the latter are all unitarily conjugate. That is, if V and V' are cartan subalgebras, there is a group element U such that $V' = \{\mathrm{ad}_U X \mid X \in V\}$. So one only has to consider conjugacy inside the standard Cartan subalgebra. (In the noncompact case, the eigenvectors correspond to representatives from any conjugacy class, which may be several in the same irreducible representation). This is enough to give the spectrum, and—in the discrete case—the full spectral resolution. In the continuous case, one still needs to find the spectral density and, from it, the S-matrix; see Kerimov [155, 156].

7 Quantum field theory and quantum statistical mechanics

This chapter discusses the basics of quantum field theory and its macroscopic consequences, given by quantum statistical mechanics. For simplicity, we concentrate on 4-dimensional relativistic quantum field theory in Minkowski space-time, and ignore additional problems associated with unsolved problems in gravity.

On the formal (uninterpreted) level, the formal core of quantum physics is valid for both quantum mechanics and quantum field theory. But since the algebra of quantities considered is different, there are two essential differences between quantum mechanics and quantum field theory:

First, in place of position and momentum operators of finitely many particles in quantum mechanics, one has in quantum field theory operators for fields. Each field ϕ has a space-time dependence that satisfy Galilei or Poincaré invariance and causal commutation relations. Thus, each field provides an infinitude of uncertain quantitites. More precisely, since from a rigorous point of view, field operators $\phi(x,t)$ at spatial position x and time t are distribution-valued operators, the basic quantities in quantum field theory are **smeared fields**, local space-time integrals

$$\phi(f) = \int_\Omega f(x,t)^T \phi(x,t)\, dx\, dt$$

over local patches Ω in space-time, where f is a smooth **test function** (for example, a Gaussian), and multipoint generalizations of these.

Second, unlike in quantum mechanics, position in quantum field theory is not an operator but a parameter, hence has no associated uncertainty. The uncertainty is instead in the quantities described by the details of the test functions f, associated with real field measurements. The fact that in quantum field theory, position is a classical parameter, whereas in quantum mechanics it is an uncertain quantity strongly, affects the relation between quantum field theory and reality.

In Section 7.1, we define the meaning of the general notion of a field for the abstract setting from Section 6.6 and comments on relativistic quantum field theory at finite times, a usually much neglected topic. Section 7.2 then shows how coherent spaces may be used to define relativistic quantum field theories. Nothing more than basic definitions and properties are given; details will be given elsewhere. Section 7.3 discusses smeared field expectations and pair correlation functions, which are among the important computable properties in quantum field theory. They encode most of what is of experimental relevance in quantum field theory.

All statistical mechanics is based on the concept of coarse-graining, introduced in Section 7.5. Statistical mechanics proper starts with the discussion of Gibbs states (Section 7.6) and the statistical thermodynamics of equilibrium and nonequilibrium (Section 7.7), leading—for example—to the Navier–Stokes equations of fluid mechanics.

https://doi.org/10.1515/9783110667387-007

Other ways of coarse-graining lead to quantum-classical models (Sections 7.8 and 7.9), generating among others the Born–Oppenheimer approximation widely used in quantum chemistry. Section 7.10 shows in which sense classical statistical mechanics is a special case of quantum statistical mechanics.

7.1 Fields and their dynamics

The description of dynamics in current relativistic quantum field theory textbooks is a delicate subject. In many such books, physical meaning is given only to scattering processes, that is, the behavior at asymptotic times $t \to \pm\infty$, whose statistical properties are expressed in terms of time-ordered correlation functions. Textbooks commonly restrict their attention to the calculation of the low-order contributions to the scattering amplitudes and how these are renormalized to give finite results. Questions about the quantum field dynamics at finite times are not discussed, since the dynamics is deeply buried under the formal difficulties of the renormalization process needed to make relativistic quantum field theory work. Sometimes they are even claimed to be impossible to address![1] However, on the level of rigor customary in theoretical physics, quantum field dynamics at finite time is actually well-defined in terms of the so-called closed time path (CTP) approach; see, for example, Calzetta & Hu [59].

In the general Lie algebraic framework of Sections 6.5–6.6, a **field** is an element ϕ of the space of $\mathbb{L}$-valued distributions $\mathbb{L} \otimes S(M,V)^*$, satisfying

$$\frac{\partial}{\partial x_\nu}\phi(x) = p_\nu \angle \phi(x). \tag{7.1}$$

Here $S(M,V)^*$ is the dual of the Schwartz space $S(M,V)$ of rapidly decaying smooth functions on M with values in V (or the space of $\mathbb{L}$-valued sections of a corresponding fiber bundle with generic fiber V). Thus, the **smeared fields**

$$\phi(f) := \int_M dx f(x)\phi(x),$$

defined for arbitrary test functions $f \in S(M,V)$, provide quantities in $\mathbb{L}$. The properties of fields are primarily functions of the distribution-valued q-expectations

$$\phi_{\mathrm{cl}}(x) = \langle \phi(x) \rangle_0$$

1 For example, Scharf [261] writes in his introduction to Chapter 2: "The more one thinks about this situation, the more one is led to the conclusion that one should not insist on a detailed description of the system in time. From the physical point of view, this is not so surprising, because in contrast to non-relativistic quantum mechanics, the time behavior of a relativistic system with creation and annihilation of particles is unobservable. Essentially only scattering experiments are possible, therefore we retreat to scattering theory. One learns modesty in field theory."

of fields and the distribution-valued **Greens functions**

$$W(x,y) = \langle \phi(x), \phi(y) \rangle_0$$

of field products at some fixed spacetime origin 0. After smearing with test functions, these distributions produce proper numbers. A comparison of (7.1) with (6.11) shows that

$$\phi(x+z) = \phi(x)^{T_z}.$$

As a consequence, field expectations from different spacetime views satisfy

$$\langle \phi(x) \rangle_z = \langle \phi(x+h) \rangle_{z-h},$$

showing that the choice of a fixed origin is inessential; a change of origin only amounts to a spacetime translation.

We also see that for any quantity $A \in \mathbb{L}$, the definition

$$A(x) := A^{T_x} \quad \text{for } x \in M$$

defines a field. These fields are more regular than the fields occurring in relativistic quantum field theory; the latter are proper distributions.

Since the traditional Schrödinger picture breaks manifest Poincaré invariance, relativistic quantum field theory is almost always treated in the Heisenberg picture. The Heisenberg dynamics on the fields is given by

$$\frac{\partial}{\partial x_\nu} \phi(x) = p_\nu \angle \phi(x),$$

where p_ν is the νth operator component of the **4-momentum** vector p, defined as the generator of the translations of the Poincaré group. In particular, the physical Hamiltonian $H = cp_0$, where c is the speed of light, is obtained after the construction of the N-point functions (q-expectations of fields and q-correlations) as the operator generating the time shift of the fields.

In place of the time-dependent q-expectations in the nonrelativistic setting discussed before, we consider q-expectations dependent on a space-time location x in Minkowski space, and require that the resulting space-time dependent q-expectations $\langle A \rangle_x$ satisfy the covariant Ehrenfest equation (6.13). One easily concludes that for arbitrary space-time points $x, y, z, w, \ldots$,

$$\langle \phi(z) \rangle_x = \langle \phi(z+x-y) \rangle_y,$$

generalizing the nonrelativistic (2.11), and

$$\langle \phi(z)\phi(w)\cdots \rangle_x = \langle \phi(z+x-y)\phi(w+x-y)\cdots \rangle_y.$$

Therefore, the complete spacetime-dependence of q-expectations, and in particular their dynamics, is determined by the q-expectations at any particular fixed space-time position.

From the covariant Ehrenfest picture, we may now deduce a **covariant Schrödinger picture** by writing

$$\langle A\rangle_x := \operatorname{Tr}\rho(x)A$$

with a space-time dependent density operator ρ. Then (6.13) becomes the **covariant von Neumann equation**

$$i\hbar\frac{\partial}{\partial x_\nu}\rho(x) = [p_\nu, \rho(x)].$$

The rank is preserved. Hence, if ρ has rank 1, this is equivalent with writing $\rho(x) = \psi(x)\psi(x)^*$, where $\psi(x)$ satisfies the **covariant Schrödinger equation**

$$i\hbar\frac{\partial}{\partial x_\nu}\psi(x) = p_\nu\psi(x).$$

After rescaling by the speed of light, we may define a 4-dimensional time $t_\nu = xv/c$ and a 4-dimensional energy $H_\nu = cp_\nu$ that turn this equation into a perfect covariant analogue

$$i\hbar\frac{\partial}{\partial t_\nu}\psi(x) = H_\nu\psi(x).$$

of the nonrelativistic Schrödinger equation. Thus it seems that in quantum field theory, time and energy—not as usually said, space and momentum—have become 4-dimensional! Ordinary time and energy are just the components t_0 and E_0.

7.2 Coherent spaces for quantum field theory

The techniques of geometric quantization do not easily extend (except on a case by case basis) to the quantization of infinite-dimensional manifolds, which would be necessary for modeling quantum field theories. However, the coherent space approach extends to quantum field theory. The coherent manifolds are now infinite-dimensional, and their topology is more technical to cope with than in the finite-dimensional case. The process of second quantization is such an example of quantization of infinitely many degrees of freedom. Thus, second quantized calculations become tractable via infinite-dimensional coherent spaces. For example, the calculus of creation and annihilation operators was developed in Neumaier & Ghaani Farashahi [212] in terms of Klauder spaces, giving simple proofs of many standard results on calculations in Fock spaces.

The groups that can be most easily quantized are infinite-dimensional analogues of the symplectic, orthogonal and (for fixed particle number) unitary groups, Kac–Moody groups, some related groups, and their abelian extensions. For example, the homogeneous quadratic expressions in finitely many creation and annihilation operators form a symplectic Lie algebra in the Boson case (CCR), and an orthogonal Lie algebra in the Fermion case (CAR); see Zhang et al. [316].

This explains why knowing the representation theory of these groups (in the form of implications for their coherent spaces) is important.

Free quantum field theories are essentially the large N limit of the finite case. Large N amounts to discretizing configuration space or momentum space, keeping only N degrees of freedom. This is the basis of lattice methods. The thermodynamic limit $N \to \infty$ creates convergence problems—one has to struggle to avoid undefined expressions producing the infamous "infinities". The correct way to do this requires some functional analysis and introduces cocycles (that, for finite N, are trivial, and hence can be avoided). For actual calculations (by computer), one needs everything as explicitly as possible, and coherent spaces yield explicit formulas for the things of interest.

In quantum field theory, one needs to take the limit analytically rather than numerically, and a key problem is to decide when these limits exist, and whether one can find them constructively enough to get useful conclusions. Using these formulas allows one to replace the usual long-winded calculations with operators in the second-quantized formalism by fairly short arguments.

Some work on infinite-dimensional versions is available. In particular, for Bosons, one needs the metaplectic representation of infinite-dimensional symplectic groups, constructed in terms of Gaussians, and for Fermions one needs the spin representation of infinite-dimensional orthogonal groups, constructed in terms of Pfaffians. The paper Gracia-Bondía & Várilly [105], though not very readable, contains lots of details (but not in terms of coherent spaces), and shows that the representation theory is enough to settle the case of **quantum electrodynamics** (**QED**) in an external field. This is easier than full QED since the field equations are linear. The mathematical challenge is the extension to nonlinear fields. (In [105], applications to the nonlinear case are promised for a follow-up paper, but such a paper never appeared.) In QED proper, asymptotic electrons are infraparticles (cf. item 9 in Section 14.3) rather than standard massive particles. Though we do not have a conventional Fock space, the asymptotic structure of QED is reasonably well understood. See, for example, the work by Herdegen [124, 125] and Kapec et al. [152].

Measures in infinite dimensions. For the quantization of infinite-dimensional manifolds, the Hilbert space is traditionally constructed as a space of integrable functions with respect to a measure on the manifold. Constructing the right measure is difficult since there is no translation invariant measure that could take the place of Lebesgue measure in finite dimensions. Thus, geometric quantization becomes an ad hoc procedure in each particular case. On the other hand, the coherent space approach

generalizes, without severe problems, to infinite dimensions. Second quantization thus appears as the theory of highest weight representations of infinite-dimensional Lie groups, or rather its coherent space version, which is somewhat simpler to manage. That it works in 2 dimensional spacetime is illustrated by the success of conformal field theory, which has a rigorous mathematical description in terms of highest weight representations of the Virasoro group.

This relation between the quantum world and the classical is important in quantum field theory when it comes to the explanation of perturbatively inaccessible phenomena, such as particle states corresponding to solitons, or tunneling effects related to instantons; see JACKIW [145]. However, his explanations are mathematically vague. A coherent space setting makes this mathematically rigorous, at least in the semiclassical approximation.

Form factors. Form factors appear as coefficients of operators in the algebra of quadratics in the defining fields satisfying a conservation law (that is, vanishing divergence). They determine the possible interactions with gauge fields. A good description of form factors requires a detailed knowledge of the causal irreducible representations of the Poincaré group. See, for example, WEINBERG [294, 295], and KLINK [162]; the results there are not manifestly covariant. The coherent space approach can be used to give nicer, manifestly covariant formulas. The form factors of a theory are needed for a subsequent analysis of spectral properties, such as the Lamb shift in QED.

Causal coherent manifolds. A **spacetime** is a smooth real manifold M with a Lie group $\mathbb{G}(M)$ of distinguished diffeomorphisms, called **spacetime symmetries** and a symmetric, irreflexive **causality relation** $\times$ on M preserved by $\mathbb{G}(M)$. We say that two sections j, k of a vector bundle over M are **causally independent** and write this as $j \times k$ if

$$x \times y \quad \text{for } x \in \operatorname{Supp} j,\ y \in \operatorname{Supp} k.$$

Here $\operatorname{Supp} j$ denotes the support of the function j.

A **causal coherent manifold** over a spacetime M is a coherent manifold Z with the following properties:

(i) The points of Z form a real vector space of smooth sections of a vector bundle over M.
(ii) The symmetries in $\mathbb{G}(M)$ act as unitary coherent maps.
(iii) The coherent product satisfies the following causality conditions:

$$K(j, j') = 1 \quad \text{if } j \times j' \text{ or } j \,||\, j' \tag{7.2}$$

$$K(j + k, j' + k) = K(j, j') \quad \text{if } j \times k \times j'. \tag{7.3}$$

Examples of important spacetimes include:

(i) **Minkowski spacetime** $M = \mathbb{R}^{1\times d}$ with a Lorentzian inner product of signature $(+^1, -^d)$ and $x \times y$ iff $(x - y)^2 < 0$. Here d is the number of spatial dimensions; most often $d \in \{1, 3\}$. $\mathbb{G}(M)$ is the **Poincaré group** $ISO(1, d)$.

(ii) **Euclidean spacetime** M with $x \times y$ iff $x \neq y$. Two Euclidean cases are of particular interest:
(iii) For **Euclidean field theory**, $M = \mathbb{R}^4$ and $\mathbb{G}(M)$ is the group $ISO(4)$ of Euclidean motions.
(iv) For **chiral conformal field theory**, M is the unit circle and $\mathbb{G}(M)$ is the **Virasoro group**. Its center acts trivially on M, but not necessarily on bundles over M.

To give examples of a causal coherent manifold, we mention that from any Hermitian quantum field ϕ of a relativistic quantum field theory satisfying the **Wightman axioms**, for which the smeared fields $\phi(j)$ (with suitable smooth real test functions j) are self-adjoint operators, and any associated state $\langle\cdot\rangle$; the definition

$$K(j,j') := \langle e^{-i\phi(j)} e^{i\phi(j')} \rangle$$

defines a causal coherent manifold.

There are many known classes of relativistic quantum field theories satisfying these properties in 2 and 3 spacetime dimensions. Under additional conditions one can conversely derive from a causal coherent manifold the Wightman axioms for an associated quantum field theory. In 4 spacetime dimensions, only free and quasifree examples satisfying the Wightman axioms are known. The question of the existence of interacting relativistic quantum field theories in 4 spacetime dimensions is completely open.

Many tools from finite-dimensional analysis, in particular the Lebesgue integral, Liouville measure, and averaging over compact sets must be replaced by more unwieldy constructs, and limits need much more careful considerations. Lack of heeding this would lead to the familiar ultraviolet (UV) divergences and infrared (IR) divergences of conventional quantum field theories. The IR and UV divergences go away if the mathematically rigorous and correct considerations are applied. This can be seen in quantum field theories in 2 and 3 spacetime dimensions.

It is an open problem how to achieve the same in interacting quantum field theories in the most important case, 4-dimensional spacetime. There it is only known how to avoid the UV divergences, using careful distribution splitting techniques in the context of causal perturbation theory (Scharf [261]). However, this approach only gives constructions for asymptotic series and misses the nonperturbative contributions needed for a rigorous definition of interacting relativistic quantum field theories in 4 spacetime dimensions.

7.3 Observability in quantum field theory

The most directly observable features of a macroscopic system modeled by quantum field theory are the q-expectations of smeared versions of the most important quantum fields, integrated over cells of macroscopic or mesoscopic size. Indeed, statistical mechanics allows one to derive for the q-expectations of the fields the equations of state

of equilibrium thermodynamics for cells of macroscopic size in thermal equilibrium, and the hydromechanical equations for cells of mesoscopic size in local equilibrium. Both are known to yield excellent macroscopic descriptions of matter.

For macroscopic systems, one must necessarily use a coarse-grained description in terms of a limited number of parameters. In the quantum field theory of macroscopic objects, any averaging necessary for applying the law of large numbers is already done inside the definition of the macroscopic (that is, smeared) operator to be measured. As shown in statistical mechanics, this is sufficient to guarantee very small uncertainties of macroscopic q-observables. Thus, one does not need an additional averaging in terms of multiple experiments on similarly prepared copies of the system. This is the deeper reason why quantum field theory can make accurate predictions for single observations of macroscopic systems.

The q-expectations $\langle\phi(x)\rangle$ of fields are distributions that produce the—in principle approximately measurable—numbers $\langle\phi(f)\rangle$ when integrated over sufficiently smooth localized test functions f. Certain **q-correlations**, q-expectation of a product of operators at pairwise distinct points, are also measurable in principle by probing the state with external fields in linear or nonlinear response theory. See, for example, Hänggi & Thomas [113] for classical correlations and Davies & Spohn [65] for quantum correlations.

Scattering experiments provide further observable information, about time-ordered multipoint correlations of these fields. The related S-matrices also appear in microlocal kinetic descriptions of dilute macroscopic matter at the level of the Boltzmann equation or the Kadanoff–Baym equations. These are derived from the q-expectations of products of fields at two different space-time points. (The kinetics of the Boltzmann equation derived from the particle picture has long been replaced by more accurate Kadanoff–Baym equations derived from field theory.)

2-point correlations in quantum field theory are effectively classical observables; indeed, in kinetic theory they appear as the classical variables of the Kadanoff-Baym equations, approximate dynamical equations for the 2-point functions. After a Wigner transform and some further approximation (averaging over small cells in phase space), these turn into the classical variables of the Boltzmann equation. After integration over momenta and some further approximation (averaging over small cells in position space), these turn into the classical variables of the Navier-Stokes equation, hydromechanic equations that describe the behavior of macroscopic fluids. For macroscopic solids, one can use similar approximations to arrive at the equations of elasticity theory. The most detailed classical level, the Kadanoff-Baym equations, still contains the unsmeared ensemble means of field products.

There is nothing in quantum field theory apart from q-expectations of the fields and q-correlations. The quantities accessible to an observer are those q-expectations and q-correlations, whose arguments are restricted to the observer's world tube. More precisely, what we can observe is contained in the least oscillating contributions to

these q-expectations and q-correlations. The spatial and temporal high-frequency part is unobservable due to the limited resolution of our instruments.

All macroscopic objects are objects describable by hydromechanics and elasticity theory; so their classical variables have the same interpretation. Thus, the quantum-mechanical ensemble averages are classical variables. Moreover, because of the law of large numbers, $\langle f(x)\rangle \approx f(\langle x\rangle)$ for any sufficiently smooth function f of not too many variables. (These caveats are needed because high dimensions and highly nonlinear functions do not behave so well under the law of large numbers.) In particular, we get from Ehrenfest's theorem (2.9) the standard classical Hamiltonian equations of motion for macroscopic objects.

Statistical mechanics shows that the uncertainties in the macroscopically relevant smeared fields scale with the inverse square root of the cell volume. This means that integrals over $\langle\Phi(\mathbf{x},t)\rangle$ are macroscopically meaningful only to an accuracy of order $V^{-1/2}$, where V is the volume occupied by the mesoscopic cell containing $\mathbf{x}$, assumed to be homogeneous and in local equilibrium. This is the standard assumption for deriving from first principles hydrodynamical equations and the like. It means that the interpretation of a field gets more fuzzy as one reduces the amount of coarse graining—until at some point the local equilibrium hypothesis is no longer valid.

Everything deduced in quantum field theory about macroscopic properties of matter in local equilibrium or dilute matter in the kinetic regime follows, and one has a completely self-consistent setting. The transition to classicality is automatic and needs no deep investigations: The classical situation is simply the limit of a huge number of degrees of freedom, where the law of large numbers discussed in Section 3.6 reduces the uncertainty to a level below measurement accuracy.

7.4 Currents

We consider the example of modeling an electric current. From a quantum field theoretical point of view, an electric current consists (in the situation to be discussed here) of the motion of the electron field in a wire at room temperature. In the field theoretic description, one has a fermionic effective electron field $\psi(x)$, and the theoretically exact current density is described by the distribution-valued q-expectation

$$J^\mu(x) = \operatorname{Tr}\rho j^\mu(x)$$

determined by the current operator

$$j^\mu(x) = -e : \overline{\psi}(x)\gamma^\mu\psi(x) : .$$

Here the colons denote normal ordering, and ρ is the density operator describing (in the Heisenberg picture) the exact state of the quantum field. Denoting by k the **Boltz-**

mann constant, we define the **entropy operator**[2] of a positive definite state by

$$S := -k \log \rho, \tag{7.4}$$

so that $\rho = e^{-S/k}$. This expression for S is exact by definition; in actual numerical work, approximations are introduced when one replaces the exact S by a numerically tractable expression. In particular, close to thermal equilibrium, it is well-established empirical knowledge that we have

$$S \approx (H + PV - \mu N)/T;$$

equality defines exact equilibrium. We can substitute this (or a more accurate non-equilibrium) approximation into the defining formula for $J(x)$ to compute a numerical approximation.

Ignoring reading uncertainties, one measures—in practice—always an electric current of the form

$$I(t) = \int dz h_t(z) \cdot J(x+z) \tag{7.5}$$

flowing at time t through a cross section of the galvanometer. Here $h_t(z)$ is a smearing function that is negligible for z larger than the size of the current-sensitive part of the galvanometer. The precise h can be found by calibration.

The smearing is needed for mathematical reasons to turn the distribution-valued current into an observable numerical vector, and for physical reasons, since a meter measuring the current is insensitive to very high spatial or temporal frequencies. This smearing has nothing to do with coarse-graining: It is also needed in already coarse-grained classical field theories. For example, in hydromechanics, the Navier–Stokes equations generally have only weak (distributional) solutions that make numerical sense only after smearing.

Thus, in field theory, the quantum situation is not very different from the classical situation. Nowhere in our exposition was any statistical argument used; the trace (which in traditional statistical mechanics gets a statistical interpretation) is simply a calculational device for managing the q-expectations.

For other currents everything is analogous. On the basis of relativistic and nonrelativistic scattering theory (Haag [110], Ruelle [253], Sandhas [259]), it can be shown that one can canonically associate with every bound state of a Poincaré invariant relativistic or Galilei invariant nonrelativistic quantum field theory a distinguished effective 4-vector current operator. This allows one to represent all asymptotic scattering phenomena using currents in place of particles, giving a finite time quantum field picture of these processes.

2 This name seems appropriate since $\langle S \rangle = \operatorname{Tr} \rho S = -k \operatorname{Tr} \rho \log \rho$ agrees with the traditional entropy.

7.5 Coarse-graining

The same system can be studied at different levels of resolution. When we model a dynamical system classically at high enough resolution, it must be modeled stochastically since the quantum uncertainties must be taken into account. But at a lower resolution, one can often neglect the stochastic part and the system becomes deterministic. If it were not so, we could not use any deterministic model at all in physics, but we often do, with excellent success.

Coarse-graining explains the gradual emergence of classicality, due to the law of large numbers to an ever increasing accuracy as the sizes of the objects grow. The quantum dynamics changes gradually into classical dynamics. The most typical path is through nonequilibrium thermodynamics (see Section 7.7). There are also intermediate stages modeled by quantum-classical dynamics (see Section 7.8); these are used in situations where the quantum regime is important for some degrees of freedom but not for others. In fact, there is a wide spectrum of models leading from full quantum models over various coarse-grained models to models with a fully classical dynamics. One typically selects from this spectrum the model that, given a desired accuracy, is most tractable computationally.

A coarse-grained model is generally determined by singling out a vector space R of **relevant quantities**, whose q-expectations are the variables in the coarse-grained model. If the coarse-grained model is sensible, one can describe a deterministic or stochastic **reduced dynamics** of these variables alone, ignoring all the other q-expectations that enter the deterministic Ehrenfest dynamics (see Section 2.2) of the detailed description of the system.

Note that the same situation in the reduced description corresponds to a multitude of situations of the detailed description. Hence, each of its realizations belongs to different values of the q-expectations in the environment, slightly causing the realizations to differ. Thus, any coarse-graining results in small prediction errors, which usually consist of neglecting certain experimentally inaccessible high-frequency effects. These uncontrollable errors are induced by the variables in the environment and introduce a stochastic element in relation to the experiment even when the coarse-grained description is deterministic.

To give a concrete example of coarse-graining, we mention Jeon & Yaffe [149], who derive the hydrodynamic equations from quantum field theory for a real scalar field with cubic and quartic self-interactions. To do so, they identify field expectations with the classical values of the field.

There are many systems of practical interest, where the most slowly varying degrees of freedom are treated classically, whereas the most rapidly oscillating ones are treated in a quantum way. The resulting quantum-classical dynamics, discussed in Section 7.8, also constitutes a form of coarse-graining. The approximation of fields (with an infinite number of degrees of freedom) by finitely many particles is also a form of coarse-graining.

In the context of coarse-graining models given in a Hamiltonian quantum framework, the **Dirac–Frenkel variational principle** (which in a coherent space setting is the coherent action principle of Section 6.1) may be profitably used for coarse-graining whenever a pure state approximation is reasonable. This principle is based on the fact that the integral

$$I(\psi) = \int \psi(t)^*(i\hbar\partial_t - H)\psi(t)\,dt = \int \big(i\hbar\psi(t)^*\dot{\psi}(t) - \psi(t)^*H\psi(t)\big)\,dt \tag{7.6}$$

is stationary iff ψ satisfies the time-dependent Schrödinger equation $i\hbar\dot{\psi}(t) = H\psi(t)$. Suppose now that a family of pure states ϕ_z (depending smoothly on a collection z of labels ranging over a coherent space) is believed to approximate the class of states realized in Nature, we may make the coarse-graining ansatz

$$\psi(t) = \phi_{z(t)}$$

and determine the time-dependent parameters $z(t)$ by finding the differential equation for the stationary points of $I(\phi_z)$, varied over all smooth functions $z(t)$. This variational principle was first used by Dirac [69] and Frenkel [91], and found numerous applications; see Kramer & Saraceno [166].

A typical phenomenon arising in coarse-grained models of detailed quantum systems involving a large environment is **decoherence**; see, e. g., Schlosshauer [264, 265]. It says that in a suitable reduced description, the density operators soon get very close to diagonal, recovering after a very short decoherence time a Koopman picture of classical mechanics. Therefore, decoherence provides, in principle (though it is rarely viewed in these terms), a reduction of the quantum physics of an open system to a highly nonlinear classical stochastic process.

Systematic projection operator or path integral techniques for coarse-graining in more general situations, given a fundamental quantum field theoretic description, can be found in Balian [20], Calzetta & Hu [59], Grabert [104], and Rau & Müller [245]. In general, once the choice of the resolution of modeling is fixed, this fixes the amount of approximation tolerable in the ansatz, and hence the necessary list of extensive quantities. What is necessary is not always easy to see, but can often be inferred from the practical success of the resulting coarse-grained model.

7.6 Gibbs states

The detailed state of a quantum system can be found with a good approximation only for fairly stationary sources of very small objects, of which sufficiently many can be prepared in essentially the same quantum state. In this case, one can calculate sufficiently many expectations by averaging over the results of multiple experiments on these objects, and use these to determine the state via some version of quantum state

tomography [309]. Except in very simple situations, the result is a mixed state described by a density operator. Mixed states are necessary also to properly discuss properties of subsystems (see Chapter 15 of the Appendix) and for the realistic modeling of dissipative quantum systems by equations of Lindblad type (Lindblad [174]). Even for the multiphoton states used to experimentally check the foundations of quantum physics, quantum opticians use density operators and not wave functions, since the latter do not provide the efficiency information required to rule out loopholes.

Although only a coarse-grained description of a macroscopic system can be explicitly known, this does not mean that the detailed state does not exist. *The existence of an exact state for large objects has always been taken as a metaphysical but unquestioned assumption.* Even in classical mechanics, it was always impossible to know the exact state of the solar system with sun, planets, asteroids, and comets treated as rigid bodies. But before the advent of quantum mechanics shattered the classical point of view, its existence was never questioned.

In quantum statistical mechanics, one only considers states that are **Gibbs states**, that is, described by density operators of the form[3]

$$\rho := e^{-S/k}, \tag{7.7}$$

where k is the Boltzmann constant and S is a self-adjoint Hermitian quantity, called the **entropy** of the system in the given state. (The traditional entropy is the uncertain value $\langle S\rangle$ of the present quantity S.) Note that a unitary transform $\rho' = U\rho U^*$ of a Gibbs state by a unitary operator U is again a Gibbs state. Indeed, the entropy of the transformed state is simply $S' = USU^*$. This shows that the notion of a Gibbs state is dynamically well-behaved; the von Neumann dynamics ensures that we get a consistent evolution of Gibbs states.

By construction, the density operator (7.7) is the complete, exact description of the state, not a coarse-grained one. However, the exact S is usually unknown, and one obtains a coarse-grained reduced description by replacing the exact S with a suitable approximate S given by a more tractable parameterized expression.

The simplest and perhaps most important case of a Gibbs state is that of an **equilibrium state** of a pure substance, defined by the formula

$$S = (H + PV - \mu N)/T,$$

3 This is only a slight restriction of generality, excluding certain more idealized states, for example, pure states. All states, including the idealized ones, are obtainable as limits of Gibbs states. This is because the positive definite density operators are dense in the set of all density operators, and every positive definite density operator is a Gibbs state. Indeed, being trace class and Hermitian, a density operator is self-adjoint, and positive definiteness implies the existence of the self-adjoint entropy operator $S = -k \log \rho$, showing that (7.7) holds. In particular, it is experimentally impossible to distinguish between a pure state and Gibbs states sufficiently close to the pure state.

where H is the Hamiltonian, V is the system volume, N a nonrelativistic number operator, and **temperature** T, **pressure** P, and **chemical potential** μ are parameters. This represents equilibrium states in the form of density operators corresponding to a **grand canonical ensemble**, $\rho = e^{-\beta(H+PV-\mu N)}$, where $\beta = 1/kT$.

A complete derivation of equilibrium thermodynamics in terms of grand canonical ensembles is given in Chapter 10 of Neumaier & Westra [214]. In this development, there is no mention of size. The latter matters only when one wants to conclude exact thermodynamic results, since then the thermodynamic limit (infinite volume limit) has to be taken into account to reduce the uncertainty to zero.

A realistic system is never exactly in equilibrium, but if it is sufficiently close to equilibrium, the entropy S is well-approximated by its equilibrium expression $(H + PV - \mu N)/T$. The residual term $H + PV - \mu N - ST$, which vanishes at equilibrium, contains the detailed information thrown away in the equilibrium approximation.

7.7 Nonequilibrium statistical mechanics

In most coarse-grained models used in statistical mechanics, the form assumed for the entropy operator S is a linear combination of **relevant quantities**, whose q-expectations define the **extensive variables** of the description. The corresponding coefficients are parameters characterizing the particular state of the reduced system; they are referred to as the **intensive variables** of the description. Extensive variables scale linearly with the size of the system (which might be mass, or volume, or another additive parameter), whereas intensive variables are invariant under a change of system size. We do *not* use the alternative convention to call "extensive" any variable that scales linearly with the system size, and "intensive" any variable that is invariant under a change of system size.

If the relevant quantities depend on continuous variables, which is the case in nonequilibrium situations, the extensive and intensive variables become fields depending on the continuum variables used to label the subsystems. For extensive variables, the integral of their field quantities over the label space gives the bulk value of the extensive quantity; thus the fields themselves have a natural interpretation as a **density**. For intensive variables, an interpretation as a density is physically meaningless; instead, they have a natural interpretation as **field strengths**, sources for **thermodynamic forces** given by their gradients.

In statistical mechanics, we distinguish four nested levels of thermal descriptions, depending on whether the system is considered to be in global, local, microlocal, or quantum equilibrium. The highest and computationally simplest level, **global equilibrium**, is concerned with macroscopic situations characterized by finitely many space- and time-independent variables.

The next level, **local equilibrium**, treats macroscopic situations in a continuum mechanical description, leading, for example, to the Navier–Stokes equations of fluid

mechanics. Here the equilibrium subsystems are labeled by the space coordinates. Therefore, the relevant variables are finitely many space- and time-dependent fields. The entropy operator S becomes time-dependent and is represented as a spatial integral

$$S(t) := \int s(t,x)\,dx$$

with a spatial entropy density $s(t,x)$. For a pure monatomic substance, the latter is in the nonrelativistic case of the form

$$s(t,x) = \big(\varepsilon(t,x) + p(t,x) - \mu(t,x)\rho(t,x)\big)/T(t,x),$$

where $\varepsilon(t,x)$ and $\rho(t,x)$ are the internal energy density and the mass density operators of a quantum field theory, whose expectations give extensive densities, and $T(t,x)$, $p(t,x)$, and $\mu(t,x)$ are intensive coefficient fields defining the local temperature, pressure, and chemical potential. (In the relativistic case, similar but more involved expressions are used, and the identification of temperature and pressure is convention-dependent.)

The next deeper level, **microlocal**[4] **equilibrium**, treats mesoscopic situations in a kinetic description, where the equilibrium subsystems are labeled by phase space coordinates. This leads, for example, to the Boltzmann equation or the Kadanoff–Baym equations. The relevant variables are now finitely many fields, depending on time, position, and momentum; see Balian [20] and Rau & Müller [245]. Now the entropy operator S is represented (in the nonrelativistic case) as a phase space integral

$$S(t) := \int s(t,x,p)\,dx\,dp$$

with a phase space entropy density $s(t,x,p)$ linearly expressed in terms of Wigner-transformed operators of a quantum field theory, whose expectations give extensive phase space densities.

The bottom level is the microscopic regime, where we must consider **quantum equilibrium**. This no longer fits a thermodynamic framework, but must be described in terms of quantum dynamical semigroups and dynamical equations of Lindblad type (Lindblad [174]).

Each description level may be considered as a special case of each more detailed description level. For example, global equilibrium is a special case of local equilibrium; the extensive variables in the single-phase global equilibrium case have constant densities.

4 The term microlocal for a phase space dependent analysis is taken from the literature on partial differential equations; see, for example, Martinez [180]. A more traditional nomenclature would call this the **kinetic regime**.

In phenomenological approaches to nonequilibrium thermodynamics, the entropy operator is written as a linear combination

$$S = \left(H - \sum_j \alpha_j X_j \right) / T,$$

of relevant extensive quantities X_j, when space is not resolved, and a corresponding density form, when space is resolved to local equilibrium. (In microlocal equilibrium, temperature T is no longer well-defined, and the linear combination is written differently.) In each case, the relevant quantities are precisely those variables that are observed to make a difference in modeling the phenomenon of interest. Table 7.1 gives typical extensive variables (S and X_j), their intensive conjugate variables (T and α_j), and their contribution (TS and $\alpha_j X_j$) to the Euler equation

$$H = TS + \sum_j \alpha_j X_j \tag{7.8}$$

resulting from the definition of the entropy. Some of the extensive variables and their intensive conjugates are vectors or (in elasticity theory, the theory of complex fluids, and in the relativistic case) tensors; see Balian [19] for the electromagnetic field and Beris & Edwards [36], Öttinger [224] for complex fluids.

Table 7.1: Typical conjugate pairs of thermal variables and their contribution to the Euler equation. The signs are fixed by tradition. (In the gravitational term, m is the vector with components m_j, the mass of a particle of kind j, g the acceleration of gravity, and h the height.)

Extensive X_j	**Intensive α_j**	**Contribution $\alpha_j X_j$**
entropy S	temperature T	thermal, TS
particle number N_j	chemical potential μ_j	chemical, $\mu_j N_j$
conformation tensor C	relaxation force R	conformational $\sum R_{jk} C^{jk}$
strain ε^{jk}	stress σ_{jk}	elastic, $\sum \sigma_{jk} \varepsilon^{jk}$
volume V	pressure $-P$	mechanical, $-PV$
surface A_S	surface tension γ	mechanical, γA_S
length L	tension J	mechanical, JL
displacement q	force $-F$	mechanical, $-F \cdot q$
momentum p	velocity v	kinetic, $v \cdot p$
angular momentum J	angular velocity Ω	rotational, $\Omega \cdot J$
charge Q	electric potential Φ	electrical, ΦQ
polarization P	electric field strength E	electrical, $E \cdot P$
magnetization M	magnetic field strength B	magnetical, $B \cdot M$
electromagnetic field F	electromagnetic field strength $-F^s$	electromagnetic, $-\sum F^s_{\mu\nu} F^{\mu\nu}$
mass $M = m \cdot N$	gravitational potential gh	gravitational, ghM
energy-momentum U	metric g	gravitational, $\sum g_{\mu\nu} U^{\mu\nu}$

In general, which quantities need to be considered depends on the resolution with which the system is to be modeled—the higher the resolution, the larger the family of extensive quantities. Whether we describe bulk matter, surface effects, impurities, fatigue, decay, chemical reactions, or transition states, the general setting remains the same, since it is a universal approximation scheme, whereas the number of degrees of freedom increases with increasingly detailed models.

7.8 Conservative mixed quantum-classical dynamics

The Koopman representation makes classical systems look quantum. It is also possible to makes quantum systems look classical. The resulting quantum-classical dynamics has important applications.

Since the differences between classical mechanics and quantum mechanics disappear in the Ehrenfest picture in favor of the common structure of a classical Hamiltonian dynamics, we can use this framework to mix classical mechanics and quantum mechanics. The resulting **quantum-classical dynamics** is described in many places, such as in Peres & Terno [234], Kapral & Ciccotti [153], Prezhdo & Kisil [243], Prezdho [242], Breuer & Petruccione [53]. The derivation of quantum-classical dynamics from pure quantum dynamics in these papers follows (under well-understood conditions) from the principles of statistical mechanics of q-expectations and belongs to the formal core of quantum physics, as it does not depend on any measurement issues.

There are many systems of practical interest, which are treated in a hybrid quantum-classical fashion, where the most slowly varying degrees of freedom are treated classically, whereas the most rapidly oscillating ones are treated in a quantum way.

The basic equations for a large class of quantum-classical models are, in the Schrödinger picture, the **Liouville equation**

$$i\hbar\dot{\rho} = [H(p,q),\rho] \tag{7.9}$$

and the **Hamilton equations**

$$\dot{q} = \operatorname{Tr}\rho\frac{\partial}{\partial p}H(p,q), \quad \dot{p} = -\operatorname{Tr}\rho\frac{\partial}{\partial q}H(p,q). \tag{7.10}$$

Here $q = q(t)$, $p = p(t)$ are classical, time-dependent variables, not quantum operators. $H(p,q)$ is, for any fixed p,q, a linear operator on some Euclidean space $\mathbb{H}$ of smooth wave functions, and $\rho = \rho(t)$ is a time-dependent density operator on $\mathbb{H}$. The sufficiently nice functions of q-expectations

$$\langle A(p,q)\rangle = \operatorname{Tr}\rho A(p,q), \tag{7.11}$$

where A is a (p,q)-dependent operator on a nuclear space, are classical quantities forming a commutative algebra. In terms of q-expectations, we have

$$\dot{A} = \langle H \angle A \rangle,$$

and in particular

$$\dot{q} = \left\langle \frac{\partial}{\partial p} H(p,q) \right\rangle, \quad \dot{p} = -\left\langle \frac{\partial}{\partial q} H(p,q) \right\rangle,$$

This looks like the original form of the Ehrenfest equations (2.10), except that on the left-hand side, we have classical variables and no expectations. The expected energy $\langle H(p,q) \rangle$ is conserved.

The quantum-classical dynamics preserves the rank of the density ρ. In particular, if ρ has the rank 1 form

$$\rho = \psi\psi^* \tag{7.12}$$

at some time, it has at any time the form (7.12) with time-dependent ψ. The fact that ρ has trace 1 translates into the statement that the state vector ψ is normalized to $\psi^*\psi = 1$. As discussed in detail in Section 2.6, the Liouville equation (7.9) holds iff the state vector ψ, determined by (7.12) up to a phase, satisfies—for a suitable choice of the phases—the Schrödinger equation

$$i\hbar\dot{\psi} = H(p,q)\psi.$$

In terms of the state vector, q-expectations now take the familiar form

$$\langle A(p,q) \rangle = \psi^* A(p,q)\psi.$$

The quantum-classical dynamics is commonly discussed in the Schrödinger picture, but it is independent of the picture used. The equivalent Heisenberg dynamics is

$$\frac{d}{dt}A = \frac{\partial A}{\partial q}\left\langle \frac{\partial H}{\partial p} \right\rangle - \frac{\partial A}{\partial p}\left\langle \frac{\partial H}{\partial q} \right\rangle + \frac{i}{\hbar}[H,A],$$

where now $\langle\cdot\rangle$ is the fixed Heisenberg state. From this, one can immediately see that everything depends only on q-expectations by taking expectations in this equation,

$$\frac{d}{dt}\langle A \rangle = \left\langle \frac{\partial A}{\partial q} \right\rangle \left\langle \frac{\partial H}{\partial p} \right\rangle - \left\langle \frac{\partial A}{\partial p} \right\rangle \left\langle \frac{\partial H}{\partial q} \right\rangle + \left\langle \frac{i}{\hbar}[H,A] \right\rangle. \tag{7.13}$$

This is now a fully deterministic equation for q-expectations of the mixed quantum-classical model, considered in the Ehrenfest picture. This is now the most natural picture, since we still get a Hamiltonian description of the form (2.14). But now the Lie algebra is the direct product of the Lie algebra of the classical subsystem and the Lie

algebra of the quantum subsystem. This results in a nonlinear dependence on expectations. Such nonlinearities are common for reduced descriptions obtained by coarse-graining (see Section 7.6), both from a pure quantum theory or from a classical stochastic theory (in the Koopman representation discussed in Section 7.10). Since quantum-classical systems (at least as they appear in the literature) are also coarse-grained descriptions, there is nothing surprising in that the same phenomenon occurs.

In the Schrödinger picture and the Heisenberg picture, the description of a quantum-classical system looks different from that in the purely classical and purely quantum case.

New in quantum-classical systems—compared to pure quantum dynamics—is that in the Heisenberg picture, the Heisenberg state occurs explicitly in the differential equation for the dynamics. But it does not take part in the dynamics, as it should be in any good Heisenberg picture. The state dependence of the dynamics is not a problem for practical applications since the Heisenberg state is fixed anyway by the experimental setting.

7.9 Important examples of quantum-classical dynamics

There are many examples of quantum-classical dynamics of practical importance.

Probably the most important quantum-classical system is a version of the Born–Oppenheimer approximation of molecules, widely used in quantum chemistry. Here the nuclei are described in terms of classical phase space variables, whereas the electrons are described quantum mechanically by means of a state vector ψ in a Hilbert space of antisymmetrized electron wave functions.

A spinning relativistic electron, though having no purely classical description, can be modeled quantum-classically by classical phase space variables p, q, and a quantum 4-component spin with Hamiltonian

$$H(p,q) = \alpha \cdot p + \beta m + eV(q) \tag{7.14}$$

is a 4×4 matrix parameterized by classical 3-vectors $p = p(t)$ and $q = q(t)$, $\rho = \rho(t)$ is a positive semidefinite 4×4 matrix of trace 1, and the trace in equation (7.11) is just the trace of a 4×4 matrix. One gets the equations (7.9) and (7.10) from Dirac's equation and Ehrenfest's theorem by an approximation involving coherent states for position and momentum. This is just a toy example; more useful field theoretic quantum-classical versions (see, for example, GÉRARD et al. [95]) lead to well-known Vlasov equations for (p, q)-dependent 4×4 densities, describing a fluid of independent classical electrons of the form (7.14).

Other examples include the Schrödinger-Poisson equations in semiconductor modeling and the quantum Boltzmann equation, with spin represented by 4×4 (or

in the nonrelativistic case 2×2) matrices parameterized by classical phase space variables. (On the other hand, the quantum-Boltzmann equation for spin zero is already a purely classical equation, since all its dynamical variables are mutually commuting.)

With even more realism, one needs to add to quantum-classical descriptions (see Section 7.8) a dissipative collision term accounting for interactions, and (7.14) is no longer adequate but needs additional stochastic terms.

7.10 Koopman's representation of classical statistical mechanics

Classical mechanics can be written in a form that looks like quantum mechanics. Such a form was worked out by Koopman [165] for classical statistical mechanics. In the special case, where one restricts the expectation mapping to be a $*$-algebra homomorphism, all uncertainties vanish, and the Koopman representation describes deterministic classical Hamiltonian mechanics.

We discuss classical statistical mechanics in terms of a commutative Euclidean $*$-algebra $\mathbb{E}$ of **random variables**, that is, Borel measurable complex-valued functions on a Hausdorff space Ω, where bounded continuous functions are strongly integrable, and the integral is given by $\int f := \int d\mu(X) f(X)$ for some distinguished measure μ. (For a rigorous treatment see Neumaier & Westra [214].) The quantities and the density operator ρ are represented by multiplication operators in some Hilbert space of functions on phase space. The classical Hamiltonian $H(p,q)$ is replaced by the **Koopman Hamiltonian**

$$\widehat{H} := \frac{\partial H(p,q)}{\partial q} i\frac{\partial}{\partial p} - \frac{\partial H(p,q)}{\partial p} i\frac{\partial}{\partial q}.$$

Then both in classical and in quantum statistical mechanics, the state is a density operator. The only difference between the classical and the quantum case is that in the former case, all operators are diagonal. In particular, the classical statistical mechanics of macroscopic matter is also described by (diagonal) Gibbs states.

As discussed in Section 2.2, functions of expectations satisfy a Hamiltonian dynamics given by a Poisson bracket. It is not difficult to show that the Koopman dynamics, resulting in this way from the Koopman Hamiltonian, exactly reproduces the classical Hamiltonian dynamics of arbitrary systems, in which the initial condition is treated stochastically. The Koopman dynamics is—like von Neumann's dynamics—strictly linear in the density matrix. But the resulting dynamics is highly nonlinear when rewritten as a classical stochastic process. This is a paradigmatic example for how nonlinearities can naturally arise from a purely linear dynamics.

Because of the Koopman representation, everything said in the following about quantum statistical mechanics applies as well to classical statistical mechanics.

8 Requirements for good foundations

In this chapter, we inquire how foundations independent of measurement could look like.

Section 8.1 briefly reviews Born's rule, the core of the traditional interpretation and its limitations (discussed in detail in Chapter 14 of the Appendix). Section 8.2 discusses the reasons why the axiomatic method made geometry and number theory self-interpreting mature theories. Section 8.3 poses the question whether quantum physics can be given a similar status as geometry, and introduces Callen's criterion as the appropriate bridge between theory and reality. Section 8.4 discusses requirements for objective properties relevant for quantum theory. Section 8.5 treats the question of how the universe may be considered as a quantum system. The final Section 8.6 gives a classical view of the qubit, known by Malus and Stokes already long before the advent of quantum mechanics. It gives a model for a successful interpretation of certain quantum phenomena in fully realistic terms.

8.1 Interpreting the formal core

> *Das Quadrat $|b_{nm}|^2$ ist gemäß unserer Grundhypothese die Wahrscheinlichkeit dafür, daß das System sich nach Ablauf der Störung im Zustand m befindet.*
> *[...] eine gewisse Wahrscheinlichkeit dafür besteht, daß das Atom im n-ten Zustand ist. Wir behaupten nun, daß als Maß dieser Zustandswahrscheinlichkeit die Größe $|c_n|^2 = |\int \psi(x,t)\psi_n^*(x)dx|^2$ zu wählen ist.*
>
> Max Born, 1927 [45, p. 171]

> *To me it must seem a mistake to permit theoretical description to be directly dependent upon acts of empirical assertions.*
>
> Albert Einstein, 1949 [80, p. 674]

In Section 2.1, we presented Axioms (A1)–(A6) defining the formal core of quantum physics. In addition to these formal axioms, one needs at least a rudimentary interpretation relating the formal part to experiments.

The following **minimal interpretation** seems to be universally accepted:

(MI) Upon measuring at times t_l $(l = 1, \dots, n)$ a vector X of q-observables with commuting components, for a large collection of independent identical (particular) systems closed for times $t < t_l$, all in the same state

$$\rho(t_l) = \rho \quad (l = 1, \dots, n)$$

(one calls such systems **identically prepared**), the measurement results are statistically consistent with independent realizations of a random vector X with measure as defined in Axiom (A5).

https://doi.org/10.1515/9783110667387-008

Note that (MI) is no longer a formal statement since it neither defines what **measuring** is, nor what **measurement results** are, and what **statistically consistent** or **independent identical system** means. Thus, (MI) has no mathematical meaning—it is not an axiom, but already part of the interpretation of formal quantum physics.

(MI) relates the axioms to a nonphysical entity, the social conventions of the community of physicists. The terms "measuring", "measurement results", and "statistically consistent" already have informal meaning in the reality as perceived by a physicist. Everything stated in Axiom (MI) is understandable by every trained physicist. Thus, statement (MI) is not an axiom for formal logical reasoning, but a bridge to informal reasoning in the traditional cultural setting that defines what a trained physicist understands by reality.

The lack of precision in statement (MI) is on purpose, since it allows the statement to be agreeable to everyone in its vagueness; different philosophical schools can easily fill it with their own understanding of the terms in a way consistent with the remainder.

Interpretational axioms necessarily have this form, since they must assume some unexplained common cultural background for perceiving reality. (This is even true in pure mathematics, since the language stating the axioms must be assumed to be common cultural background.)

(MI) is what *every* traditional interpretation I know of assumes at least implicitly in order to make contact with experiments. Indeed, all traditional interpretations I know of assume much more, but they differ a lot in what they assume beyond (MI).

Everything beyond (MI) seems to be controversial; see Schlosshauer [264] and [265, Chapter 8]. In particular, already what constitutes a measurement of X is controversial. (For example, reading a pointer, different readers may read marginally different results. What is the true pointer reading? Does passing a beam splitter or a polarization filter count as a measurement?)

Thus, measurement should not figure at all in the foundations of physics. The case for this was already vividly made by Bell [32]. Note that Born's probability interpretation was originally phrased in a factual language that did not involve the notion of measurement: *"The square $|b_{nm}|^2$ is according to our basic hypothesis the probability for the system to be in state m after completion of the interaction [...] there is a certain probability that the atom is in the nth state. We now claim that as measure for this probability of state, one must choose the quantity $|c_n|^2 = |\int \psi(x,t)\psi_n^*(x)dx|^2$."*

An analysis given in Chapter 14 of the Appendix shows that actual measurement practice is in conflict with the traditional foundations, due to a far too idealized view of measurement. More specifically, the unrestricted rule (MI) implies the universal Born rule discussed in Section 14.2 of the Appendix. Therefore, the critique of the universal Born rule given in Section 14.3 also applies (MI). Thus, (MI) can be valid only for a restricted class of measurements.

8.2 The interpretation of mature theories

In speaking of the copy and the original we may assume that words are akin to the matter which they describe; when they relate to the lasting and permanent and intelligible, they ought to be lasting and unalterable, and, as far as their nature allows, irrefutable and immovable-nothing less. But when they express only the copy or likeness and not the eternal things themselves, they need only be likely and analogous to the real words. As being is to becoming, so is truth to belief.

Plato, Timaeus, ca. 360 BC [239, par.29]

6. Mathematische Behandlung der Axiome der Physik.
Durch die Untersuchungen über die Grundlagen der Geometrie wird uns die Aufgabe nahe gelegt, nach diesem Vorbilde diejenigen physikalischen Disciplinen axiomatisch zu behandeln, in denen schon heute die Mathematik eine hervorragende Rolle spielt; dies sind in erster Linie die Wahrscheinlichkeitsrechnung und die Mechanik. [...] Soll das Vorbild der Geometrie für die Behandlung der physikalischen Axiome maßgebend sein, so werden wir versuchen, zunächst durch eine geringe Anzahl von Axiomen eine möglichst allgemeine Klasse physikalischer Vorgänge zu umfassen. [...] 'Das Endziel', so hat Weierstrass einmal gesagt, 'welches man stets im Auge behalten muß, besteht darin, daß man über die Fundamente der Wissenschaft ein sicheres Urteil zu erlangen suche'.

David Hilbert, 1900 [132]

If we conceive Euclidean geometry as the science of the possibilities of the relative placing of actual rigid bodies and accordingly interpret it as a physical science, and do not abstract from its original empirical content, the logical parallelism of geometry and theoretical physics is complete.

Albert Einstein, 1934 [79, p. 165]

The ordinary language, (spiced with technical jargon for the sake of conciseness) is thus inseparably united, in a good theory, with whatever mathematical apparatus is necessary to deal with the quantitative aspects. It is only too true that, isolated from their physical context, the mathematical equations are meaningless: but if the theory is any good, the physical meaning which can be attached to them is unique.

Leon Rosenfeld, 1957 [250, p. 41]

I feel induced to contradict emphatically an opinion that Professor L. Rosenfeld has recently uttered in a meeting at Bristol, to the effect that a mathematically fully developed, good and self-consistent physical theory carries its interpretation in itself, there can be no question of changing the latter, of shuffling about the concepts and formulae.

Erwin Schrödinger, 1958 [270, p. 170]

A great physical theory is not mature until it has been put in a precise mathematical form, and it is often only in such a mature form that it admits clear answers to conceptual problems.

Arthur Wightman, 1976 [303, p. 158]

In spite of the interpretational chaos just mentioned, there is an informal consensus on how to perform measurements in practice. Good foundations, including a good measurement theory, should be able to properly justify this informal consensus by defining additional formal concepts about what constitutes measurement. To be satisfying, these must behave within the theory just as their informal relatives with the same name behave in reality.

This informal consensus is achieved elsewhere in physics through the axiomatic method, a tradition started and promoted by David Hilbert. He wrote in 1924 the first (and very influential) textbook on mathematical physics (Courant & Hilbert [62]), created the concept of a Hilbert space, and contributed in 1928 significantly to the foundations of quantum physics (Hilbert et al. [133]).[1]

In his famous 1900 address, Hilbert [132] clarified, in the context of the sixth problem, what an axiomatization of physics should mean: *"The investigations on the foundations of geometry suggest the problem: To treat in the same manner, by means of axioms, those physical sciences in which already today mathematics plays an important part."* This is a quest for giving axioms for probability (done by Kolmogorov in 1933) and mechanics (at that time only classical mechanics, but today it means quantum mechanics) that share the clarity and interpretation independence of geometry achieved through the axiomatic method.

The purpose of a system of axioms is to separate the stuff that is problematic but peripheral from the stuff that is essential and allows rational deductions. This enables getting universal agreement about the essence of a theory, without being entangled with controversies about their interpretation.

Axioms specify in unambiguous terms all properties that are ascribed to the concepts used, whereas interpretation rules tell informally (and in detail often debatably) how these concepts are applied as models of the real world. Thus, the axioms precisely define what the theory is about, and the interpretation rules use the concepts defined by the theory and apply them informally to the intended aspects of reality.

According to Hilbert, one can study how to achieve an informal consensus about the interpretation of a mature theory by looking at the modern account of the oldest of the physical sciences, Euclidean geometry. Indeed, for Euclidean geometry, considered as a branch of physics, there is a complete consensus about how, on laboratory scales, theory and reality correspond.[2]

One first defines the corresponding calculus and names the quantities that can be calculated from quantum mechanical models (or models of the theory considered) with the appropriate names from experimental geometry. Thus, initially, a circle was

1 It is no accident that today's quantum mechanics is based on Hilbert spaces rather than wave functions and their probability interpretation!

2 Older, in today's view nonaxiomatic interpretation rules for Euclidean geometry are ambiguous and approximate, of the kind:
A point is what has no parts.
A point is an object without extension.
A point is a mark on paper.
These are three different, mutually incompatible but common interpretation rules for a point.
This sort of observations prompted Hilbert to promote the axiomatization of theories as a means for making the content of a theory as precise as possible, separating the objective substance from the controversial philosophy.

a material object with a round shape, and the mathematical circle was an abstraction of these.

The Pythagoreans (and later Descartes and Hilbert with even more precision) then developed a theory that gives a precise formal meaning to all the geometrical concepts. This is pure mathematics, today encoded in textbook linear algebra and analytic geometry. The theory and the nomenclature were *developed* with the goal of enabling this identification in a way consistent with tradition.

Starting with Plato, the theory took precedence, defining objective **truth**—the perfect, unalterable concept. What was found through experience was viewed as subjective **belief**—an approximate, imperfect realization, an analogous copy of the theoretical concept.

This was done by declaring anything in real life resembling an ideal point, line, plane, circle, et cetera, to be a point, line, plane, circle, et cetera, if and only if it can be assigned in an approximate way (determined by the heuristics of traditional measurement protocols, whatever they are) the properties that the ideal point, line, plane, circle, et cetera, has, consistent to the assumed accuracy with the deductions from the theory. If the match is not good enough, we can explore whether an improvement can be obtained by modifying measurement protocols (devising more accurate instruments or more elaborate error-reducing calculation schemes, et cetera), or by modifying the theory (to a non-Euclidean geometry, say, which uses the same concepts, but assumes slightly different properties relating them).

No significant philosophical problems are left; lucent, intuitive, and logically impeccable foundations for Euclidean geometry were established in this way. This indicates the maturity of Euclidean geometry as a scientific discipline.

The same holds for number theory defined by the Peano axioms, as the following parable (a modified version of [210]) shows:

> *Once upon a time – not long after Viete had introduced his innovative concept of variables –, Dnikeded, an ambitious student of math, was sitting in Prof. Onaep's class, then the leading capacity in the field of applied algebra. He is reading one of Onaep's exercises: "I is a rebmun. If Z is a rebmun then ZI is a rebmun. If ZI=YI then Z=Y. Never ZI=I. Every rebmun is generated in this way."*
>
> *Dnikeded has not the slightest idea what his professor is talking about. He had never heard of anything called rebmun. Determined to figure out the meaning and being already familiar with Viete's work, he plays with the statements given.*
>
> *Well, at least he knows that I is a rebmun. Setting Z=I he discovers that II is a rebmun. Setting Z=II he discovers that III is a rebmun. Setting Z=III he discovers that IIII is a rebmun. Setting Z=IIII he discovers that IIIII is a rebmun. This reminds him of counting. Each new rebmun is obtained by adding an I to the previous rebmun. The process goes on for ever....*
>
> *Remembering what he had learnt already about algebra, Dnikeded noticed that the rebmuns could be interpreted in terms of stuff he was familiar with – numbers. If he identified I with 1 then he could equate II with 2, III with 3, IIII with 4, IIIII with 5, etc. "Ah, this is a variant of the way we count*

the number of beers in the pub," he thought, "except that each 5th bar would be drawn vertically, a minor issue that doesn't really change things."

But well, there were more properties: If ZI=YI then Z=Y. "True – if my friend and I both order a beer and then have the same number of beers, we must have had the same number of beers before. Thus Onaeps theory is predictive, and things come out correctly. Let me try the next item, never ZI=I; can I falsify my interpretation?" He tries and finds no problem with it – I is too short to be of the form ZI.

Dnikeded is left with the final statement to be figured out. He thinks about what he can generate so far: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, . . . *a never ending list of numbers. But neither 0 nor fractions like 2/3. Also no negative numbers. Suddenly everything makes sense. "Ah, I finally understand. Rebmuns are nothing else than the numbers I have been familiar with since childhood, before I got interested in more advanced numbers!"*

When Dnikeded compared his solution of the exercise with that of his friend Rotnac, he noticed that the latter had another way of interpreting Onaep. He had also played with the statements in Onaep's riddle and associated it with marbles in his pocket. He linked changing Z to ZI to putting a new marble into the pocket. Starting with the empty pocket that contained no marble, he got the correspondence I=0, II=1, III=2, IIII=3, etc..

Both Dnikeded and Rotnac tried to figure out who made an error and whose interpretation was defective. But they couldn't find one. So they went to Onaep, asking for his judgment. Onaep declared both interpretations to be valid.

Indeed, the modern concept of natural numbers (based on the Peano axioms) exists in two forms, and the two different traditions have two different interpretations, depending on whether they call 0 a natural number (for example, friends of C++ and set theorists) or whether they do not (for example, friends of Matlab and everyone before Cantor).[3] The two interpretations are related by the fact that $x \to x + 1$ is an isomorphism between the two. This is analogous to the interpretation of classical mechanics, which is unique only up to the choice of an orthonormal coordinate system. In the latter case, a rigid motion provides the necessary isomorphism.

Thus, once a theory is mature, the identification with real life is done in terms of the formal, purely mathematical theory developed, giving an interpretation to the theory. In this way, physics inherits the clarity of mathematics, the art and science of precise concepts and relations.

In particular, with the Lagrangian and Hamiltonian formulations, classical mechanics has also reached the status of maturity, and hence is perceived by most physicists as clear and philosophically unproblematic.

In summary, foundations should be concise, unambiguous, and simple. The only way to get sound, unambiguous foundations of a theory of physics is to give clear, fully precise axioms for the formal, mathematical part, then describe its consequences, and

3 I belong to the second category and believe that 0 is an unnatural number since it took ages to discover 0, and many more centuries to declare it natural. I have never seen anyone count 0, 1, 2, 3....

finally, with the conceptual apparatus created by the theory (of course with lots of hindsight, arrived at through prior, less rigorous stages) to specify the conditions when it applies to reality in a more informal way, but still attempting to preserve as much clarity as possible.

Note that one cannot separate the mathematics of a theory from its physics. A mathematical theory *is* a theory of physics once its concepts agree with those of a branch of physics, and its assumptions and conclusions can be brought into correspondence with physical reality, no matter how informal (or even unspoken) the interpretation rules are.

8.3 Is quantum physics a mature theory?

> *Operationally, a system is in an equilibrium state if its properties are consistently described by thermodynamic theory.*
>
> Herbert Callen, 1975 [58, p. 15]

To define quantum physics (or any other physical theory) properly, including a logically impeccable interpretation, one should, therefore, proceed as in Euclidean geometry and classical mechanics.

One first needs to define the corresponding calculus; this has already been done in Section 2.1. Then one has to name the quantities that can be calculated from quantum mechanical models (or models of the theory considered) with the appropriate names the experimental physicists use for organizing their data.

One can then develop a theory that gives a precise formal meaning to the concepts physicists talk about. This is pure mathematics—the shut-up-and-calculate part of quantum physics.

Finally, one needs an interpretation of the theory, that is, the identification with real life. At first it seems trivial to make this precise—just state Born's rule. But Born's rule contains the notion of "measuring"—what does this mean, in precise terms? This opens a Pandora box....

Let me give a concrete example. To define what it means to measure a time interval, we cannot proceed without first having a definition of the unit of time in which to make the measurement. The official definition [221] is:

> *"The second is the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the Cesium 133 atom."*

To be able to make sense of this interpretation rule, one needs to assume a lot of detailed theory of quantum mechanics—formal mathematical theory of what can be deduced from the formal axioms for quantum mechanics (this makes no reference to reality). This is required to understand the meaning of ground state of an atom and what its hyperfine levels are. We also need rules that tell us the real life meaning of a

Cesium 133 atom—how to verify that an atom claimed to be a Cesium 133 atom actually is a Cesium 133 atom. More formal theory enters the relation between theory (the notion of a "second") and reality (what the notion means).

One also needs to assume some additional informal rules that tell how to measure transitions between two energy levels, and how to prepare a Cesium 133 atom so that the quantity described can be measured. To understand the latter, one needs more results from quantum mechanics of the formal, mathematical kind, and more informal rules that tell how these results are interpreted in an experiment. To understand these, one needs.... et cetera.

One ends up with a whole book on measurement theory instead of simple axioms for the interpretation. Whereas it is not unreasonable to have such a book, it *is* unreasonable to have a book-sized axiom system.

Fortunately, such a book-sized axiom system is not needed: All it takes to recognize a Cesium 133 atom is to verify that it behaves like the theoretical model of a Cesium 133 atom—by whatever experimentators learn how to do it. This is indeed the only criterion; if an atom does not behave like that, we conclude with certainty that it is not a Cesium 133 atom. Thus, the whole book-sized axiom system is implied by the theory together with a single criterion,

(CC) Callen's criterion: Operationally, a system is in a given state if its properties are consistently described by the theory for this state.

This generalizes the way H. B. Callen justified in the above quote phenomenological equilibrium thermodynamics in his famous textbook (Callen [58]). At first sight, this sounds like a circular definition (and indeed Callen classifies it as such). But a closer look shows there is no circularity since the formal meaning of "consistently described by thermodynamic theory" is already known. The operational definition simply moves this formal meaning from the domain of theory to the domain of reality by defining when a real system deserves the designation "is in an equilibrium state". In particular, this definition allows one to determine experimentally whether or not a system is in equilibrium.

For quantum physics, Callen's criterion asserts that we may declare anything in real life resembling an ideal photon, electron, atom, molecule, crystal, ideal gas, et cetera, to be a photon, electron, atom, molecule, crystal, ideal gas, et cetera, if and only if it can be assigned in an approximate way (determined by the heuristics of traditional measurement protocols, whatever that is) the properties that the ideal photon, electron, atom, molecule, crystal, ideal gas, et cetera, has, consistent to the assumed accuracy with the deductions from the theory.

In practice, relevant quantities and corresponding states are assigned to real life situations by well-informed judgment concerning the behavior of the equipment used. The validity of the assignment is experimentally tested by comparing experimental results with the chosen mathematical model. As postulated by Callen, the model defines the meaning of the concepts: The theory defines what an object is.

Callen's criterion is enough to find out in each single case how to approximately measure the uncertain value of a quantity of interest, though it may require considerable experimental ingenuity to do so with low uncertainty. The uncertain value $\overline{X}$ is considered informative only when its uncertainty σ_X is much less than $|\overline{X}|$. The required identification process is fairly independent of the way measurements are done, as long as they are capable to produce the required accuracy for the matching. Hence it carries no serious philosophical difficulties.

In general, any successful theory must be crafted in such a way that it actually applies to reality—otherwise the observed properties cannot match the theoretical description. On the other hand, as Callen's criterion notes, *we already need the theory to define precisely what it is that we observe*.

As a result, theoretical concepts and experimental techniques complement each other in a way that, when a theory reaches maturity, it has developed its concepts to the point where they are a good match to reality. Thus, we may say the following:

(R) Something in real life "is" an instance of the theoretical concept if it matches the theoretical description sufficiently well.

If the match between theory and observation is not good enough, we can explore whether an improvement can be obtained by modifying measurement protocols (devising more accurate instruments or more elaborate error-reducing calculation schemes, et cetera), or by upgrading the theory to another theory that uses the same concepts, but assumes slightly different properties relating them. The latter happened in the past for Newton's mechanics, which was upgraded to relativity theory.

It is not difficult to check that this holds not only in physics, but everywhere where we have clear concepts about some aspect of reality. Indeed, Rule (R) just amounts to a definition of what it means of something in real life to "be an X", where X is defined as a theoretical concept.

In general, an **interpretation** of a mature theory, that is, its identification with real life, is given *in terms of the formal theory developed* by means of Callen's criterion (CC) or its more informal rendering (R). This criterion is sufficient to precisely determine the interpretation only if the theory is indeed mature and fully reflects everything of relevance for its interpretation. It does not work if the theory—as the traditional foundation of quantum physics given by the 7 basic rules of Section 1.1—invokes in its axioms the notion of measurement without saying anything about what on the formal level counts as a measurement!

Nearly 100 years of modern quantum physics established informally that quantum theory is indeed an appropriate and complete model for the physical aspects of reality. On this basis, one should be able to study the measurement problem rigorously: One declares that a real **detector** (in the sense of a complete experimental arrangement, including the numerical postprocessing of raw results that gives the final result) performs a real **measurement** of an ideal quantity if and only if the following holds: Modeling the real detector as a macroscopic quantum system (with the properties assigned to it by statistical mechanics/thermodynamics) predicts raw measure-

ments such that, in the model, the numerical postprocessing of the raw results that gives the final result is in sufficient agreement with the value of the ideal quantity in the model.

Then measurement analysis is a scientific activity like any other, rather than a philosophical prerequisite for setting up a consistently interpreted quantum physics. Indeed, this is the way high-precision experiments are designed and analyzed in practice.

8.4 Objective properties

> *The scientist [...] appears as* **realist** *insofar as he seeks to describe a world independent of the acts of perception; as* **idealist** *insofar as he looks upon the concepts and theories as the free inventions of the human spirit (not logically derivable from what is empirically given); as* **positivist** *insofar as he considers his concepts and theories justified only to the extent to which they furnish a logical representation of relations among sensory experiences. He may even appear as* **Platonist** *or* **Pythagorean** *insofar as he considers the viewpoint of logical simplicity as an indispensable and effective tool of his research.* [original italic in bold]
>
> Albert Einstein, 1949 [80, p. 684]

> *Some people are hoping to reintroduce determinism in some way, perhaps by means of hidden variables or something like that, but it just doesn't work according to the accepted ideas. I might add that personally I still have this prejudice against indeterminacy in basic physics. I have to accept it because we cannot do anything better at the present time. It may be that in some future development we shall be able to return to determinism, but only at the expense of giving up something else, some other prejudice which we hold to very strongly at the present time.*
>
> Paul Dirac, 1972 [72, p. 7]

> [Quantum mechanics] *is fundamentally about the results of 'measurements', and therefore presupposes in addition to the 'system' (or object) a 'measurer' (or subject). [...] the theory is only approximately unambiguous, only approximately self-consistent. [...] it is interesting to speculate on the possibility that a future theory will not be intrinsically ambiguous and approximate. Such a theory could not be fundamentally about 'measurements', for that would again imply incompleteness of the system and unanalyzed interventions from outside. Rather it should again become possible to say of a system not that such and such may be observed to be so but that such and such be so. The theory would not be about 'observables' but about 'beables'. These beables [...] should, on the macroscopic level, yield an image of the everyday classical world, [...] the familiar language of everyday affairs, including laboratory procedures, in which objective properties – beables – are assigned to objects.*
>
> John Bell, 1972 [31, p. 687]

One of the basic problems with the traditional interpretations of quantum mechanics is the difficulty to specify precisely what counts as real. The physics before 1926 was explicitly about discovering and objectively describing the true, reliably repeatable properties of Nature, seen as objectively real.

After the establishment of modern quantum physics, the goal of physics can (according to the traditional interpretations of quantum mechanics) only be much more

modest, to systematically describe what physicists measure. Nonetheless, physics continues to make objective claims about reality that existed long before a physicist performed the first measurement, such as the early history of the universe, the composition of distant stars and galaxies, of which we can measure not more than tiny specks of light, the age of ancient artifacts dated by the radio carbon method. Physics also makes definite statements about the distant future of our solar system—independent of anyone being then around to measure it.

Thus, there is a fundamental discrepancy between what one part of physics claims and what the traditional interpretations of quantum mechanics allows one to claim. This discrepancy was discussed in a paper by Bell [31], where he introduced the concept of **beables** for objective properties.

To find objective properties, we first note that the problems created by quantum mechanics are absent in classical mechanics. Therefore, it seems that classical objects **exist** in the sense that they have objective **properties**. The classical regime is usually identified with macroscopic physics, where length and time scales are long enough that the classical approximation of quantum mechanics is accurate enough to be useful. This suggests that we look at the visible parts of quantum experiments.

Most experiments done to probe the foundations of quantum physics are done using optical devices. In quantum optics experiments, both sources and beams are extended macroscopic objects describable by quantum field theory and statistical mechanics. For example, a laser beam is simply a coherent state of the quantized electromagnetic field, concentrated in a neighborhood of a line segment in space.

The sources have properties independent of measurement, and the beams have properties independent of measurement. These are objects described by quantum field theory. For example, the output of a laser (before or after parametric down conversion or any other optical processing) is a laser beam, or an arrangement of highly correlated beams. These are in a well-defined state that can be probed by experiment. If this is done, they are always found to have the properties ascribed to them by the preparation procedure. One just needs sufficient time to collect the information needed for a quantum state tomography. The complete state is measurable in this way, reproducibly, to any given accuracy.

Neither the state of the laser nor of the beam is changed by one or more measurements at the end of the beam. Moreover, these states can be found to any desired accuracy by making sufficiently long and varied measurements of the beam; how this is done is discussed in quantum optics under the name of quantum tomography.

Thus, these properties exist independent of any measurement—just as the moon exists even when nobody is looking at it. They can be found through diligent measurement, just as properties of distant stars and galaxies. They behave in every qualitative respect just like classical properties of classical objects.

Therefore, we have found a class of objective properties: the densities, intensities, and correlation functions used to describe optical fields.

8.5 The universe as a quantum system

In some sense, quantum mechanics was always a theory of the whole universe. For example, the probabilistic interpretation of the wave function ψ of a single particle asserts that there is a typically positive (though almost everywhere exceedingly tiny) probability $p(\Omega) := \int_\Omega dx|\psi(x)|^2$ for finding the particle in an arbitrary open region Ω of the universe, no matter how far away from where it was prepared.

In relativistic quantum field theory, the basic fields are local objects in the sense that smeared fields $\phi(f)$ and $\phi(g)$ commute whenever f and g have spacelike separated support. Nevertheless, a quantum state specifies the q-expectations $\langle\phi(f)\rangle$ for arbitrary smooth test functions f and the higher-order moments (n-point correlation functions) $\langle\phi(f_1)\rangle \cdots \langle\phi(f_n)\rangle$ for arbitrary smooth test functions $f_1, \ldots, f_n$, hence makes statements about uncertain values at all space-time locations, that is, everywhere in the universe. Hence, relativistic quantum field theory necessarily describes a complete universe. As long as gravitation is not modeled—that is, in almost all applications except in cosmology—everything happens in a Minkowski spacetime.[4]

There is no known limit of validity of the principles of quantum physics. Therefore, good foundations for quantum physics must allow a consistent and deterministic relativistic quantum description of the universe from the smallest to the largest levels of modeling, including its classical aspects, without having to introduce any change in the formal apparatus of quantum physics.

The foundations for quantum physics should therefore be formulated in a way that they apply not only to small systems but to large systems such as our solar system, and even to the largest physical system, the whole universe. Here the **universe** is understood to be the smallest closed physical system containing us, hence—strictly speaking—the only closed system containing us, and therefore the only system to which unitary quantum physics applies without approximation. In particular, this implies that the universe is unique.

Our solar system can be approximately treated classically, and is usually treated in this way. But from the fundamental point of view of quantum field theory, it must be considered as a quantum system. The state of the solar system, when modeled by quantum fields, should completely specify what happens in any small space-time region within the solar system.

Traditional interpretations of Copenhagen flavor require that a quantum system is measured by an external classical apparatus. They cannot apply to the quantum field theory of our solar system, say, since we do not have access to an external classical apparatus for measuring this system. The astronomers doing measurements on the solar system are part of the system measured—a situation outside the Copenhagen setting.

4 Except for a few passing remarks concerning gravity, we assume in this book a flat spacetime. In a quantum field theory of gravity, we would need to replace the Galilei or Poincaré group by the diffeomorphism group of the spacetime manifold, which has its own foundational problems.

Cosmology studies the state of the universe in a very coarse (and partly conjectured) approximation, where even details at the level of galaxies are averaged over. Only for properties localized in the solar system do we have a much more detailed knowledge.

Of course, a more detailed discussion of the state of the universe should include gravitation, and hence would touch on the difficult, unsolved problem of quantum gravity. However, one needs at least a consistent interpretational framework in which to discuss these questions, without having to worry about whether the concepts used to formulate these questions mean anything in the context of a quantum system without external observers.

Knowing all the detailed properties, or finding its exact state, is already out of the question for a small macroscopic quantum system, such as a drop of water. Thus, as for a drop of water, one must be content with describing the state of the universe approximately. But, as in case of a drop of water, there is no physical reason to question the existence of the state of the whole universe, even though many of its details may remain unknown for ever.

A good interpretation should therefore be consistent with assigning a well-defined (though only superficially known) state to the whole universe, whose properties account for everything observable within the universe.

Indeed, since every property of a subsystem is also a property of the whole system, the state of the universe must be compatible with everything we have ever empirically observed in the universe. This implies that the state of the universe is highly constrained, since knowing this state amounts to having represented all physics accessible to us by the study of its subsystems. This constitutes a very stringent test of adequacy of a putative state of the universe.

What matters for a successful physics of the universe is that we can model (and then predict) those properties of the universe that are accessible to measurement at the temporal and spatial scales of human beings. Since the quantities of interest in a study of the universe as a whole are macroscopic, they have a tiny uncertainty and are well-determined even by an approximate state. For example, one should—in principle—be able to compute from a proposed model of the universe the values of the electromagnetic field in regions where we can measure it, and we should get excellent agreement with the measurements if the computations could actually be done.

Nothing, however, should depend on the existence of measurement devices, which were not available in the very far past of the universe.

8.6 A classical view of the qubit

It is commonly said that quantum mechanics originated in 1900 with Max Planck, reached its modern form with Werner Heisenberg and Erwin Schrödinger, got its cor-

rect interpretation with Max Born, and its modern mathematical formulation with Paul Dirac and John von Neumann. It is very little known that much earlier—in 1852, at a time when Planck, Heisenberg, Schrödinger, Born, Dirac, and von Neumann were not even born—George Stokes described all the modern quantum phenomena of a single qubit, explaining them in classical terms.

Stokes' description of a qubit is couched in the language of optics—polarized light was the only quantum system that, at that time, was both accessible to experiment and quantitatively understood. Stokes' classical observables are the functions of the components of the coherence matrix, the optical analogue of the density operator of a qubit.

The transformation behavior of rays of completely polarized light was first described in 1809 by Etienne-Louis Malus [177] (who coined the name "polarization"); that of partially polarized light in 1852 by George Stokes [279]. This section gives a modern description of the core of this work by Malus and Stokes.

We shall see that Stokes' description of a polarized quasimonochromatic beam of classical light behaves exactly like a modern quantum bit.

A ray (quasimonochromatic beam) of polarized light of fixed frequency is characterized by a state, described equivalently by a real **Stokes vector**

$$S = (S_0, S_1, S_2, S_3)^T = \begin{pmatrix} S_0 \\ \mathbf{S} \end{pmatrix}$$

with

$$S_0 \geq |\mathbf{S}| = \sqrt{S_1^2 + S_2^2 + S_3^2},$$

or by a **coherence matrix**, a complex positive semidefinite 2×2 matrix ρ. These are related by

$$\rho = \frac{1}{2}(S_0 + \mathbf{S} \cdot \boldsymbol{\sigma}) = \frac{1}{2}\begin{pmatrix} S_0 + S_3 & S_1 - iS_2 \\ S_1 + iS_2 & S_0 - S_3 \end{pmatrix},$$

where $\boldsymbol{\sigma}$ is the vector of Pauli matrices. $\operatorname{Tr}\rho = S_0$ is the **intensity** of the beam. $p = |\mathbf{S}|/S_0 \in [0,1]$ is the **degree of polarization**. Note the slight difference to density matrices, where the trace is required to be one.

A linear, nonmixing (not depolarizing) instrument (for example a polarizer or phase rotator) is characterized by a complex 2×2 **Jones matrix** T. The instrument transforms an ingoing beam in the state ρ into an outgoing beam in the state $\rho' = T\rho T^*$. The intensity of a beam after passing the instrument is $S_0' = \operatorname{Tr}\rho' = \operatorname{Tr} T\rho T^* = \operatorname{Tr}\rho T^* T$. If the instrument is lossless, the intensities of the ingoing and the outgoing beam are identical. This is the case if and only if the Jones matrix T is unitary.

Since $\det\rho = (S_0^2 - S_3^2) - (S_1^2 + S_2)^2 = S_0^2 - \mathbf{S}^2$, the fully polarized case $p = 1$, that is, $S_0 = |\mathbf{S}|$, is equivalent to $\det\rho = 0$, hence holds iff the rank of ρ is 0 or 1. In this case, the coherence matrix can be written in the form $\rho = \psi\psi^*$, with a state vector ψ

determined up to a phase. Thus, precisely the pure states are fully polarized. In this case, the intensity of the beam is

$$S_0 = \langle 1 \rangle = |\psi|^2 = \psi^* \psi.$$

A **polarizer** has $T = \phi\phi^*$, where $|\phi|^2 = 1$. It reduces the intensity to

$$S_0' = \langle T^* T \rangle = |\phi^* \psi|^2.$$

This is **Malus' law**.

An instrument with Jones matrix T transforms a beam in the pure state ψ into a beam in the pure state $\psi' = T\psi$. Passage through inhomogeneous media can be modeled by means of many slices consisting of very thin instruments with Jones matrices $T(t)$ close to the identity. If $\psi(t)$ denotes the pure state at time t, then $\psi(t + \Delta t) = T(t)\psi(t)$, so that for small Δt (the time needed to pass through one slice),

$$\frac{d}{dt}\psi(t) = \frac{\psi(t + \Delta t) - \psi(t)}{\Delta t} + O(\Delta t) = \frac{(T(t) - 1)}{\Delta t}\psi(t) + O(\Delta t).$$

In a continuum limit $\Delta t \to 0$, we obtain the time-dependent **Schrödinger equation**

$$i\hbar \frac{d}{dt}\psi(t) = H(t)\psi(t),$$

where (note that $T(t)$ depends on Δt)

$$H(t) = \lim_{\Delta t \to 0} i\hbar \frac{T(t) - 1}{\Delta t}$$

plays the role of a time-dependent Hamiltonian. Note that in the lossless case, $T(t)$ is unitary, hence $H(t)$ is Hermitian.

A linear, mixing (depolarizing) instrument transforms ρ instead into a sum of several terms of the form $T\rho T^*$. It is therefore described by a real 4×4 **Mueller matrix** (Perez & Ossikovski [237]) acting on the Stokes vector. Equivalently, it is described by a completely positive linear map on the space of 2×2 matrices, acting on the polarization matrix.

Thus, we see that a polarized quasimonochromatic beam of classical light behaves exactly like a modern quantum bit. We might say that classical optics is just the quantum physics of a single qubit passing through a medium!

Indeed, the 1852 paper by Stokes [279] described all the modern quantum phenomena for qubits, explained in classical terms. In particular, the following:

- Splitting fully polarized beams into two such beams with different, but orthogonal polarization corresponds to writing a wave function as superposition of preferred basis vectors.
- Mixed states are defined (in his paragraph 9) as arising from "groups of independent polarized streams" and give rise to partially polarized beams.

- The coherence matrix is represented by Stokes with four real parameters, in today's terms comprising the Stokes vector.
- Stokes asserts (in his paragraph 16) the impossibility of recovering from a mixture of several distinct pure states any information about these states beyond what is encoded in the Stokes vector (equivalently, the coherence matrix).
- The latter can be linearly decomposed in many essentially distinct ways into a sum of pure states, but all these decompositions are optically indistinguishable, hence have no physical meaning.

In its modern formulation via Maxwell's equations, classical partially polarized light (as described by Stokes) already requires the stochastic form of these equations, featuring—just like the full quantum description—field expectations and correlation functions; see Mandel & Wolf [178]. The coherence matrices turn into simple matrix-valued field correlation functions.

The only difference to the modern description is that the microscopic view is missing. For faint light, photodetection leads to discrete detection events—even in models with an external *classical* electromagnetic field; see the discussion in Section 12.4. The trace of ρ is the intensity of the beam, and the rate of detection events is proportional to it. After normalization to unit intensity, ρ becomes the density operator corresponding to a single detection event (aka photon).

This is a simple instance of the transition from a beam (classical optics or quantum field) description to a single particle (quantum mechanical) description.

It took 75 years after Stokes until the qubit made its next appearance in the literature, in a much less comprehensive way. In 1927, Weyl [300, pp. 8–9] discusses qubits in the guise of an ensemble ("Schwarm") of spinning electrons. Instead of the language of Stokes, the description uses the paradoxical language still in use today, where the meaning of everything must be redefined to give at least the appearance of making sense.

9 The thermal interpretation of quantum physics

If you wish to learn from the theoretical physicist anything about the methods which he uses, I would give you the following piece of advice: Don't listen to his words, examine his achievements. For to the discoverer in that field, the constructions of his imagination appear so necessary and so natural that he is apt to treat them not as the creations of his thoughts but as given realities.

Albert Einstein, 1934 [79, p. 163]

In this chapter, we define a new interpretation of quantum physics, called the **thermal interpretation**. The thermal interpretation of quantum physics (including quantum mechanics, quantum field theory, quantum statistical mechanics, and their applications) allows a consistent and deterministic relativistic quantum description of the universe from the smallest to the largest levels of modeling, including its classical aspects, without having to introduce any change in the formal apparatus of quantum physics.

The goal of the thermal interpretation is to provide foundations that match all requirements from Chapter 8. The thermal interpretation agrees with how one interprets measurements in statistical thermodynamics, the macroscopic part of quantum physics, derived via statistical mechanics. Extrapolating from the macroscopic case, the thermal interpretation considers the functions of the state (or of the parameters characterizing a state from a particular family of states) as the objective properties of what really exists. Some of these are accessible to experiment—namely the expectation values of quantities that have a small uncertainty and vary sufficiently slowly in time and space. Because of the law of large numbers, all thermodynamic variables are in this category. By its very construction, the thermal interpretation naturally matches the classical, in essence, thermodynamical properties of our quantum world.

9.1 A reinterpretation of the tradition

Expectation values are all that is required to establish empirical equivalence with orthodox quantum mechanics.

Hall, Deckert, and Wiseman, 2014 [114, p. 5]

Quite early in the history of modern quantum physics, Ehrenfest [76] found the Ehrenfest equation[1] (2.9), a clean and exact relation between the dynamical laws of classical mechanics and quantum mechanics phrased solely in terms of q-expectation. The implications for the interpretation of quantum physics seem to have gone unnoticed in the literature. Instead, tradition placed an unreasonable notion of quantum measurement at the very basis of quantum physics.

1 It is a pity that Ehrenfest did not develop this equation to the point, where it would have amounted to an interpretation of quantum physics. This could have avoided a lot of the subsequent confusion.

https://doi.org/10.1515/9783110667387-009

As we have seen in Part I, the formal part of quantum physics is mainly concerned with q-expectations and their approximation in concrete applications. For much of the applications of the formal part, nothing else is needed: The emergence of spectroscopic information in resonance situations (Section 2.4), the thermodynamics of equilibrium and nonequilibrium processes (Sections 7.6 and 7.7), the emergence of classical properties for quantities with negligible uncertainty (Section 2.3)—all this only involves the identification of certain q-expectations with classical observables or interaction terms in classical dynamics. As the classical view of the qubit (Section 8.6), known to Stokes long before the advent of quantum mechanics, shows, even the properties of streams of particles can be fully captured by q-expectations with the classical meaning of a beam intensity. In this light, sources and beams seem much more real than particles. Should not then the former, not the latter, be the real players in solid foundations?

On the level of the theory, probability and statistics play no role at all—not even in quantum statistical mechanics, where only the basic intuition is colored by statistical ideas. Probability and statistics make their appearance only when discussing the interface between theory and experiment, where their use is dictated by handling noisy open systems using multiple measurements and the law of large numbers. This is in complete agreement with how probability and statistics appear in classical physics.

This strongly suggests that q-expectations should be regarded as objective properties of quantum systems, not only in the above case, where their interpretation is time-honored, but in general, even in situations where there are seeming discrepancies, such as in experiments involving a multitude of single, experimentally distinguishable events.

The thermal interpretation takes this point of view and declares all q-expectations, and whatever can be computed from them as objective properties of a quantum system. The originally statistical machinery of q-expectations is used in an abstract way, stripped from its original meaning, in the same way as, in functional analysis or quantum mechanics, the notion of a vector is used for functions that have only formal properties in common with the vectors that gave this today abstract notion its name. This is justified by remembering that the statistical notion of uncertainty is commonly used (and recommended by the normative agencies; see Section 3.1) also for uncertainty of conceptual rather than statistical origin. Thus, the abstract use of statistical mathematics for nonstatistical purposes precedes its use in the thermal interpretation.

In the thermal interpretation, the true properties of a quantum system, approximately revealed by a measurement, are the q-expectations rather than the eigenvalues. Its link of the formal core to experiment is therefore not the eigenvalue link of the traditional interpretations, but a generalization of the link between quantum physics and thermodynamics known from statistical mechanics. This is emphasized through the attribute “thermal” in the name of the interpretation.

Remarkably, this change in the view of the nature of the true value of an imperfect measurement result brings about many benefits that make it worthwhile to revise the foundations of quantum physics on this point. The rewards are manifold; see the list of advantages given in Section 9.4.

On the other hand, there is also a price to pay: After nearly a century of conditioning to the opposite convention specified in Born's rule, this change of interpretation seems at first sight radical and very counterintuitive. One needs to rethink everything along the lines of the new interpretation, and make sure that the old problems are really solved and no new problems are introduced. This is the purpose of the following chapters.

9.2 The thermal interpretation

> *Die richtige Frage mußte also lauten: Kann man in der Quantenmechanik eine Situation darstellen, in der sich ein Elektron ungefähr – das heißt mit einer gewissen Ungenauigkeit – an einem gegebenen Ort befindet und dabei ungefähr – das heißt wieder mit einer gewissen Ungenauigkeit – eine vorgegebene Geschwindigkeit besitzt, und kann man diese Ungenauigkeiten so gering machen, daß man nicht in Schwierigkeiten mit dem Experiment gerät?*
>
> Werner Heisenberg, 1972 [122, pp. 77f]

Quantum theory describes both **single quantum systems** and **populations** of quantum systems in maximal generality in terms of a Lie $*$-algebra $\mathbb{L}$ of **quantities**, and a **state**, a positive definite bilinear functional on $\mathbb{L}$, defining the **q-expectations** of the system and their **uncertainties**. In the most typical case, $\mathbb{L}$ is the Lie algebra of a $*$-algebra $\mathbb{E}$ of linear operators on a Euclidean space, and a state is a positive linear functional on $\mathbb{E}$. The **objective properties** of a quantum system are given by its q-expectations, their uncertainties, and everything computable from these.

Beyond a certain accuracy, the uncertainty in the position of macroscopic objects, such as the sun, a city, a house, a tire, an apple, a cloud, or a water wavelet is a conceptual uncertainty impossible to resolve by measurement. The thermal interpretation asserts that all quantum uncertainty is an uncertainty of the same conceptual kind as the uncertainty of the position of an apple.

In particular, the thermal interpretation answers Heisenberg's question, *"Can quantum mechanics represent the fact that an electron finds itself approximately in a given place and that it moves approximately with a given velocity, and can we make these approximations so close that they do not cause experimental difficulties?"* by giving electrons objective but uncertain paths, making them move along an extended **world tube** rather than an infinitesimally thin classical world line.[2]

2 Even a pointlike quantum object is extended. Since it has a 3-component position vector $\mathbf{q}$, its extension is determined by the computable position uncertainty $\sigma_{\mathbf{q}} = \sqrt{\sigma^2_{\mathbf{q}_1} + \sigma^2_{\mathbf{q}_2} + \sigma^2_{\mathbf{q}_3}}$, where $\mathbf{q}$ is the

> *[...] auf die objektive Beschreibbarkeit der individuellen* Makro-*Systeme (Beschreibung des 'Realzustandes') nicht verzichtet werden kann ohne dass das physikalische Weltbild gewissermassen sich in einen Nebel auflöst. Schliesslich ist die Auffassung wohl unvermeidbar, dass die Physik nach einer Realbeschreibung des Einzel-Systems streben muss. Die Natur als Ganzes kann eben nur als individuelles (einmalig existierendes) System gedacht werden und nicht als eine 'System-Gesamtheit'.*
>
> Albert Einstein, 1953 [81, p. 40]

The **universe** (*"die Natur als Ganzes"*) is an isolated quantum system, the only isolated quantum system containing us.

> *According to the present quantum mechanics, the probability interpretation, the interpretation which was championed by Bohr, is the correct one. But still, Einstein did have a point. He believed that, as he put it, the good God does not play with dice. He believed that basically physics should be of a deterministic character.*
> *And, I think it might turn out that ultimately Einstein will be proved right, because the present form of quantum mechanics should not be considered as the final form. [...] And I think that it is quite likely that at some future time we may get an improved quantum mechanics in which there will be a return to determinism and which will, therefore, justify the Einstein point of view.*
>
> Paul Dirac, 1975 [73, p. 10]

As a quantum system, the universe is exactly modeled by a unitary dynamics with a (only crudely known) time-independent Hamiltonian. This dynamics is theoretically exact and deterministic, given by the linear **Ehrenfest equations** for q-expectations.

> *In a statistical description of nature only expectation values or correlations are observable.*
>
> Christof Wetterich, 1997 [299, p. 2678]

Quantum fields are treated as the basic entities characterizing the universe. The q-expectations of quantum fields (densities and currents) describe local features. Correlation functions, that is, q-expectations of products of quantum fields at different spacetime points account for objective **nonlocal properties** of the world. The concept of a **particle** appears only in effective descriptions appropriate under suitable conditions. In particular, objective properties of a **beam** consistently represent the elusive properties of unobserved particles traditionally said to travel along the beam.

> *One must ignore or down-grade the full information in the environmental variables to see dissipation appearing in the dynamics of the open system. [...] Coarse graining can be the causal truncation of a correlation hierarchy, the averaging of the higher modes, the "integrating out" of the fluctuation fields, or the tracing of a density matrix (discarding phase information). [...]*
> *While the dynamics of the combined system made up of a subsystem and its environment is unitarity, and its entropy remains constant in time, when certain coarse graining is introduced in the environment, the subsystem turns into an open system, and the entropy of this open system (con-*

position vector of the object. In spacetime, the uncertain positions $\langle \mathbf{q} \rangle_t$ for a range of times t trace out an uncertain world line. The quantum object can be visualized as moving along a fuzzy world tube, a nested family of world tubes formed by the union of the intervals $[\langle \mathbf{q} \rangle_t - \kappa\sigma_{\mathbf{q}}(t), \langle \mathbf{q} \rangle_t + \kappa\sigma_{\mathbf{q}}(t)]$, for reasonable κ of order one.

structed from the reduced density matrix by tracing out the environmental variables) increases in time. In this open system dynamics, the effect of the coarse-grained environment on the subsystem leads to dissipation and irreversibility in its dynamics.

Esteban Calzetta and Bei-Lok Hu, 2008 [59, p. 9]

All other quantum systems are proper **subsystems**, defined by identifying the subspace of quantities relevant for describing the subsystem. If considered without their **environment** (the remainder of the universe), proper subsystems are **open systems**, modeled approximately by a stochastic, dissipative, and often nonlinear dynamics. The exact dynamics of a quantum system is induced by the dynamics of the universe, of which it is a subsystem, and becomes independent of the environment in an approximate **effective description**, obtained through an approximation process called **coarse-graining** that tracks explicitly only the relevant quantities of the subsystem. The **stochasticity** reflects the influence of high-frequency modes of the unmodeled environment; the **dissipation** reflects the loss of information permanently leaving the system through unmodeled degrees of freedom. Often, the stochasticity or the dissipation may be neglected.

The state of every system is determined in the thermal interpretation by the state of the universe through taking the appropriate partial trace. Therefore essentially every state is mixed. Even the states that we usually regard as pure are in fact only approximately pure.[3] In the thermal interpretation, a mixed state is therefore not a sign of uncertainty about the state. If there is uncertainty about a state, it just means having an inaccurate state instead of the true one, just as in the case of uncertainty about the precise number of atoms in a piece of matter. Such an uncertainty about the state is not part of quantum theory itself but of the art of applying it to real life situations. Quantum theory is about what is objectively true, and not about which knowledge human beings have about what is objectively true.

On some limited time scale, open systems may be treated as isolated if they interact sufficiently weakly with the environment, so that both stochasticity and dissipation can be neglected. Then they are also modeled by a unitary dynamics, with a possibly time-dependent Hamiltonian. For isolated quantum systems in a pure state,[4] the dynamics is equivalent to the Schrödinger equation for the state vector describing the pure state.

Operationally, a system is in an equilibrium state if its properties are consistently described by thermodynamic theory.

Herbert Callen, 1975 [58, p. 15]

3 For example, it is impossible to determine exactly the direction of a magnetic field in a Stern-Gerlach experiment, but it would be needed to prepare an exact spin up particle.

4 Only tiny isolated quantum systems can be prepared in a pure state: Only for these, experimental procedures are known, which ensure that, relative to a given description level, theory, observation, and previous calibration experiments imply the objective properties characterizing a state.

The interpretation of quantum theory, that is, its identification with real life, is given in terms of the formal theory developed by means of **Callen's criterion** discussed in Section 8.3. Thus, operationally, a quantum system is in a given state if its properties are consistently described by the theory for this state.

The thermal interpretation differentiates between objective properties of quantum systems—which the systems possess according to a scientific model, independent of whether these properties are known or even knowable, and **experiments** consisting of a sequence of **measurements**—which are the scientist's approximate way of checking such properties and validating the corresponding models. **Experimental physics** is about how to do the measurements, and under which conditions which measurements are how accurate.

Experimentors know or assume on the basis of past experience, claims of manufacturers, et cetera, that certain materials or machines reliably produce states that, to a satisfactory degree for the purpose of the experiment or application, depend only on variables that are accounted for in our theory, and that are, to a satisfactory degree, either fixed or controllable. The nominal state of a system can be checked and, if necessary, corrected by **calibration**, using appropriate measurements that reveal the parameters characterizing the state. Similarly, all claims that certain materials or machines reliably produce certain states, and that certain instruments measure certain properties, can be checked and, if necessary, corrected by calibration. Once the assumptions about the whole equipment are made consistent with the theory, based on Callen's criterion, the calibrated equipment can be used to prepare and measure in agreement with the theory.

> *One is almost tempted to assert that the usual interpretation in terms of sharp eigenvalues is 'wrong', because it cannot be consistently maintained, while the interpretation in terms of expectation values is 'right', because it can be consistently maintained.*
>
> John Klauder, 1997 [160, p. 6]

When performing on a quantum system a measurement of a quantity A with a physical meaning, one gets an approximation for its value. The thermal interpretation treats the measured value as an approximation not of an eigenvalue of A but of the q-expectation of A, the formal expectation value defined as the trace of the product of A with a density operator describing the state of the system. The approximation error is of the order of the uncertainty σ_A. By the Chebyshev inequality (3.4) (cf. also Theorem 10.6.1 below), this postulate is essentially implied by—and hence more cautious than—the traditional postulate that the measured value is an eigenvalue, obtained with the probability given by Born's rule.

This novel postulate of the thermal interpretation remains valid in all cases where the traditional postulates apply since it is essentially a weaker version of the latter. It avoids a number of problems of Born's rule (collected in Section 14.3).

Thus, in the thermal interpretation, there is no longer a direct relationship between the objective properties of a system measured and the eigenvalues of some of its quantities. This constitutes the major departure from tradition—a most significant step away from the conventional understanding, justified in Section 10.6 below. It is the innovation that is responsible for the features that allow the thermal interpretation to solve the measurement problem.

> *Wenn wir aus jenem mathematischen Schema physikalische Resultate ableiten wollen, so müssen wir den quantentheoretischen Größen, also den Matrizen (oder 'Tensoren' im mehrdimensionalen Raum) Zahlen zuordnen. [...] Man kann also sagen: Jeder quantentheoretischen Größe oder Matrix läßt sich eine Zahl, die ihren 'Wert' angibt, mit einem bestimmten wahrscheinlichen Fehler zuordnen; der wahrscheinliche Fehler hängt vom Koordinatensystem ab; für jede quantentheoretische Größe gibt es je ein Koordinatensystem, in dem der wahrscheinliche Fehler für diese Größe verschwindet. Ein bestimmtes Experiment kann also niemals für alle quantentheoretischen Größen genaue Auskunft geben.*
>
> Werner Heisenberg, 1927 [117, pp. 181f]

Measurement is a complex derived process, in which objective properties of a measured system are correlated with objective properties of a measurement device, in such a way that more or less reliable information about the measured system can be read off from the measurement device. Measurement results are usually inaccurate approximations of what they measure. Thus, measurements reveal partial, often very noisy, information about the measured system.

The deterministic dynamics of the complete collection of q-expectations constructible from quantum fields, when restricted to the set of measurable ones, gives rise to all the stochastic features observed in practice. In particular, the probabilistic, discrete nature of certain microscopic results (such as detector clicks, individual spots on a screen, or particle tracks) arises from bistability in the effective models describing the measurement process, in a similar way as randomness, discreteness and hysteresis arise from deterministic classical dynamical systems (for example, a roulette wheel or a power switch). From this and the deterministic rules, Born's statistical interpretation follows in the limited range where it applies. (These limitations are discussed in Section 14.3 of the Appendix.)

> *Within the framework of statistical quantum theory there is no such thing as a complete description of the individual system. [...]*
> *If it should be possible to move forward to a complete description, it is likely that the laws would represent relations among all the conceptual elements of this description which, per se, have nothing to do with statistics.*
>
> Albert Einstein, 1949 [80, p. 671]

Many q-expectations can be approximately measured by reproducible single measurements of macroscopic quantities, or by sample means in a large number of observations on similarly prepared microscopic systems. If a large population of independent

but similarly prepared quantum systems is measured, the uncertainty of the individual results is reduced by the law of large numbers by a factor approximately equal to the square root of the population size. In particular, many q-probabilities can be approximately measured by determining the relative frequencies of corresponding events associated with a large number of independent, similarly prepared systems.

Statistical interpretations of quantum mechanics are thus recovered in the limit of large population size.

> *The concept of observation is in so far arbitrary as it depends upon which objects are included in the system to be observed. Ultimately every observation can of course be reduced to our sense perceptions. The circumstance, however, that in interpreting observations use has always to be made of theoretical notions, entails that for every particular case it is a question of convenience at what point the concept of observation involving the quantum postulate with its inherent 'irrationality' is brought in.*
>
> Niels Bohr, 1927 [39, p. 580]

For describing a quantum system, **classical physics** is appropriate whenever all its relevant quantities (that is, those appearing in an effective classical dynamical description) have a tiny uncertainty only, so that the Ehrenfest equation for these quantities can be approximated by the corresponding classical equation. With an appropriate choice of relevant quantities, macroscopic systems behave classically in this sense. Often, only some of the relevant quantities allow such a classical description, resulting in a mixed quantum-classical system with a hybrid quantum-classical dynamics.

The **Copenhagen interpretation** of quantum mechanics may be regarded as the special case, where the measurement process is described as a stochastic quantum-classical system in the form of a piecewise deterministic process, and the selection of classically described relevant variables defines the Heisenberg cut.

> *If one considers statistical mechanics as a form of statistical inference rather than as a physical theory, it is found that the usual computational rules, starting with the determination of the partition function, are an immediate consequence of the maximum-entropy principle. In the resulting 'subjective statistical mechanics', the usual rules are thus justified independently of any physical argument, and in particular independently of experimental verification; whether or not the results agree with experiment, they still represent the best estimates that could have been made on the basis of the information available.*
>
> Edwin Jaynes, 1977 [146, Abstract]

As everywhere in science, subjective aspects enter an otherwise objective description in different ways: (i) the delineation of a system under study from its environment (related to the Heisenberg cut in the Copenhagen interpretation); (ii) the choice of a description level, specifying the set of relevant quantities for modeling a system at a desired resolution; (iii) the lack of knowledge about the exact details of a model, and the consequent need for approximation; (iv) the quality of the approximation schemes used; (v) the quality of the data used for the estimation of the parameters of a model.

Knowledge interpretations of quantum mechanics, where a state says nothing objective about the systems modeled, but is only about the subjective knowledge of these systems, may be regarded as possible conceptual accounts for the rational management of observer knowledge about the (in these interpretations unmodeled) objective properties, represented by the q-expectations of the relevant quantities of measured quantum systems and corresponding measurement devices.

frustra fit per plura quod potest fieri per pauciora

William of Ockham, 1323 [223]

The existence of multilocal q-expectations implies that a composite system is more than its parts. The deterministic Ehrenfest dynamics of the collection of all q-expectations couples local q-expectations (that is, idealized pointer readings) to multilocal q-expectations. This accounts for the nonclassical correlations observed in long-distance entanglement.

Hidden variable interpretations of quantum mechanics introduce additional degrees of freedom into the description of a quantum system to obtain a deterministic description of quantum dynamics. Due to a well-known result by Bell [30] and its experimental verification by Aspect [17], such a deterministic description must necessarily (ignoring potential loopholes) involve nonlocal features. The thermal interpretation identifies the objective nonlocal properties and the deterministic dynamics, without having to introduce hidden variables.

Nous tenons la mécanique des quanta pour une théorie complète, dont les hypothèses fondamentales physiques et mathématiques ne sont plus susceptibles de modification.

Max Born and Werner Heisenberg, 1927 [47, p. 178]

Quantum physics determines in a huge number of cases the statistics of experiments with phenomenal success. However, it currently leaves most details about the outcomes of experiments undetermined—each single outcome of a test for the state of a qubit, the onsets of individual clicks in a counter, and all details of observed fluctuations in continuous time experiments. Indeed most of the stuff that is actually observed is left undetermined. Only their gross statistics is determined.

But only in the traditional interpretations. The thermal interpretation predicts the outcomes of experiments individually from the state of the universe, in terms of the quantum formalism alone without additional variables. It accounts for each single outcome, and for all details of the fluctuations. Only our limited knowledge of the state of the universe forces us to statistical considerations.

In view of Ockham's razor [223, 135], that we should opt for the most economic model explaining a phenomenon of interest, this suggests that, with the thermal interpretation, quantum mechanics is indeed a complete theory for which the fundamental physical and mathematical hypotheses are no longer susceptible of modification—as anticipated by Born and Heisenberg many years ago.

9.3 The interpretation of quantum-classical systems

A consistent interpretation of quantum-classical systems must be in terms of concepts that have identical form in classical and in quantum mechanics; otherwise, there are inevitable conflicts. This is impossible in the traditional statistical interpretation; there are several theorems in the literature documenting this [49, 258]. The reason is that q-expectations in the quantum-classical dynamics cannot be interpreted as the result of averaging over many realizations; they must be properties of the single quantum system, for which the dynamics is supposed to hold.

New in quantum-classical systems—compared to pure quantum dynamics—is that in the Heisenberg picture, the Heisenberg state appears explicitly in the differential equation for the dynamics—though, as it should be in any good Heisenberg picture, it does not take part in the dynamics. This makes an important difference in the interpretation of the theory. In contrast to the pure quantum case, there is now a difference between averaging results of two experiments ρ_1, ρ_2 and the results of a single experiment ρ given by $(\rho_1 + \rho_2)/2$. That, in ordinary quantum theory, the two are indistinguishable in their statistical properties is a coincidental consequence of the linearity of the Schrödinger equation, and the resulting state independence of the Heisenberg equation; it does no longer hold in effective quantum theories, where nonlinearities appear due to a reduced description. Since quantum-classical systems (at least as they appear in the literature) also are reduced descriptions, there is nothing surprising in that the same phenomenon occurs.

Because the dynamics depends on the Heisenberg state, calculating results by splitting a density at time $t = 0$ into a mixture of pure states no longer makes sense. One gets different evolutions of the operators in different pure states, and there is no reason why their combination should, at the end, give the correct dynamics of the original density. (And indeed, this will usually fail.) This splitting is already artificial in pure quantum mechanics since there is no natural way to tell of which pure states a mixed state is composed of. But there the splitting happened to be valid and useful as a calculational tool since the dynamics in the Heisenberg picture is state independent. In the quantum-classical case, not even this is possible, so the quantum-classical equations have no sensible interpretation in terms of mixing pure cases into an ensemble.

Thus, the quantum-classical setting cannot be consistently interpreted in the traditional interpretations, where q-expectations have only a statistical meaning.

On the other hand, the thermal interpretation can cope successfully with this challenge since q-expectations are irreducible objects describing a single quantum system, not stochastic entities that make sense only under repetition. Therefore, in the thermal interpretation, the quantum-classical setting is very natural. It gives meaning to a theory that contains the classical and the quantum case as two special cases of the same conceptual framework. In this framework, one can therefore discuss things con-

sistently that lead to puzzles if interpreted either on a pure classical or on a pure quantum basis, or in some ill-defined in-between limbo.

It is—in principle—conceivable (though not desirable from the point of view of simplicity) that the most fundamental description of Nature is truly quantum-classical and not purely quantum. In the absence of an interpretation with a consistent quantum-classical setting, this would have been unacceptable, but apart from elegance, there are no longer fundamental reasons that would forbid it. In particular, this could be a possible resolution of the problem of quantum gravity.

9.4 Advantages of the thermal interpretation

> *If the relations are known to us, what does it matter if we think it convenient to replace one image by another?*
>
> Henri Poincaré, 1902 [241, p. 161]

The conventions embodied in the thermal interpretation have, compared to tradition, several direct or indirect advantages.

The thermal interpretation

- allows a consistent quantum description of the universe and its subsystems, from the smallest to the largest levels of modeling, including its classical aspects;
- acknowledges that there is only one world;
- has no split between classical and quantum mechanics—the former emerges naturally as the macroscopic limit of the latter;
- is about both real systems and idealized systems, at every level of idealization;
- applies both to single quantum objects (like a quantum dot, a neutron star, or the universe) and to statistical populations;
- allows one to make definite statements about each single quantum system, no matter how large or small it is;
- is by design compatible with the classical ontology of ordinary thermodynamics;
- satisfies the principles of locality and Poincaré invariance, as required for relativistic quantum field theory;
- is compatible with relativistic causality for extended objects;
- is description-dependent but observer-independent, hence free from unspecified subjective elements;
- correctly reflects the actual practice of quantum physics, especially regarding its macroscopic implications;
- provides foundations that are easily stated and motivated since they are essentially the foundations used everywhere for uncertainty quantification;
- uses no concepts beyond what is taught in every quantum physics course;
- requires at the levels of the postulates apart from definitions no technical mathematics—no spectral theorem, no notion of eigenvalue, no probability theory;

- gives a natural, realistic meaning to the standard formalism of quantum mechanics and quantum field theory in a single world;
- preserves the agreement of quantum theory with the experimental record, without introducing changes in the kinematics or the dynamics of the established theory;
- gives a fair account of the interpretational differences between quantum mechanics and quantum field theory;
- involves no philosophically problematic steps—it eliminates from the foundations the philosophically problematic notions of probability and measurement;
- paints a deterministic picture of quantum physics in which God does not play dice—it only seems so to us mortals because of our limited resolution capacity, and since we have access to a limited part of the universe only;
- explains how the unmodeled environment influences the results enough to cause all randomness in quantum physics;
- explains the emergence of probabilities from the linear, deterministic dynamics of quantum states or quantum observables;
- derives Born's rule for scattering and in the limit of ideal measurements;[5]
- is independent of the measurement problem—which becomes a precise problem in statistical mechanics rather than a fuzzy and problematic notion in the foundations;
- solves the measurement problem, makes quantum mechanics much less mysterious, and makes it much less different from classical mechanics;
- gives at any time simultaneously idealized but uncertain values to position and momentum of distinguishable particles, eliminating the spooky nature of the traditional quantum ontologies;
- explains the peculiar features of the Copenhagen interpretation (lacking realism between measurements) and the statistical interpretation (lacking realism for the single case) in the microscopic world, where the latter apply;
- explains the collapse as emerging in the coarse-grained approximation of non-isolated subsystems;
- explains what physicists actually do rather than what they say;
- explains what people actually use in the applications (as contrasted to work on the foundations themselves), even when they pay lipservice to another interpretation.

Thus, the thermal interpretation satisfactorily resolves the traditional stumbling blocks for a clear description of the relation between the quantum formalism and experimental reality. As a consequence, the thermal interpretation leads to a gain in clarity of thought. This results in saving efforts otherwise spent in the contemplation

5 But in general, only part of Born's rule holds exactly: Whenever a quantity A with zero uncertainty is measured exactly, its value is an eigenvalue of A.

of aspects arising in traditional interpretations, which from the point of view of the thermal interpretation appear as meaningless or irrelevant side problems.

The following list points to the quantum phenomena treated in this book on the basis of the thermal interpretation, some extensively, others only shortly.

- black body radiation (Section 12.3);
- photodetection (Section 12.1);
- the Compton effect (Section 13.2);
- particle tracks (Section 12.2);
- particle impinging on a screen (Section 13.1);
- particle decay (Section 13.2);
- the qubit (Sections 8.6 and 10.3–10.4);
- motion of heavy particles (Section 2.3);
- spectroscopy (Section 2.4);
- chemical and nuclear reactions (Section 13.2);
- quantum chemistry in the Born–Oppenheimer approximation (Section 7.9);
- the Stern–Gerlach experiment (Sections 13.3–13.4);
- tests for Bell inequalities (Section 13.6).

9.5 Open problems

In terms of thermal interpretation, the measurement problem turns from a philosophical riddle into a scientific problem in the domain of quantum statistical mechanics, namely how the quantum dynamics correlates macroscopic readings from an instrument with properties of the state of a measured microscopic system. The developments in this book show that the thermal interpretation resolves the measurement problem at least in a qualitative way. However, to rightly claim a quantitative solution, a number of detailed questions need to be investigated. Unlike traditional interpretations, the thermal interpretation is an interpretation of quantum physics that is, in principle, refutable by theoretical arguments.

This section presents a number of open problems:

- The measurement principle (MP) (from Section 11.1) demands that any instrument for measuring a quantity A has an uncertainty $\Delta a \geq \sigma_A$. How to prove this from the statistical mechanics of measurement models is an open problem.
- The derivation of a piecewise deterministic stochastic process (PDP) by Breuer & Petruccione [54] (see Section 11.7) suggests that, under the standard assumptions that go into the traditional derivations in classical statistical mechanics, collapse in a single observed system—in the modern POVM version (see Section 11.5) of the corresponding von Neumann postulate for quantum dynamics – is generally derivable from the unitary dynamics of a bigger system. It would

be desirable to have for this a direct argument, not dependent on a statistical approach.[6]

- It should be possible to show in quantitative detail how position loses its parameter status and becomes uncertain when going from the relativistic quantum field description of a beam to a corresponding quantum mechanical description of a sequence of particles moving along the beam.
- It should be possible to explain from the dynamics of the universe the statistical features of scattering processes and the temporal instability of unobserved superpositions of pure states, as caused by the neglect of the environment.
- The thermal interpretation might have applications for understanding, controlling, and ultimately designing detailed dissipation properties of microscopic systems.

6 Added in proof: A general, direct and rigorous argument for the POVM version of Born's rule will be given in a forthcoming paper "A derivation of Born's rule" by the author. Born's rule itself cannot be derived, because of its limitations spelled out in Section 14.3.

10 Measurement

Die vorher scheinbar unlösbaren Paradoxien der Quantentheorie beruhten alle darauf, daß man diese mit jeder Beobachtung notwendig verbundene Störung vernachlässigt hatte.

Werner Heisenberg, 1929 [119, p. 495]

In this chapter, we discuss measurement from a fundamental point of view, where all physical systems are treated as subsystems of the universe. All measurement outcomes are functions of q-expectations of the detector, hence are deterministically predicted by the state of the universe, though not only by the state of the subsystem consisting of the measured object and the detector.

We consider in detail how the thermal interpretation explains the emergence of binary responses of a measurement device when coupled with the simplest quantum object, a qubit, with probabilities given by the diagonal entries of the reduced density matrix of the prepared qubit.

10.1 Objective properties and their measurement

The **objective properties** of a quantum system are given by its q-expectations, their uncertainties, and everything computable from these. A **statement** is a {`true`, `false`}-valued property. This gives a clear formal meaning to the notion of existence, an **ontology**. In the thermal interpretation, something about a quantum system is said to **exist**, to be **real**, and to be **objective**—three ways of expressing the same—if and only if it can be expressed solely in terms of objective properties of the system as just defined. This formal ontology is defined on the level of the model specifying the quantum system. It gives a clear formal meaning to the notion of existence. Whether something that exists in this model sense also exists in Nature depends on how faithful the model is to the corresponding aspect of Nature.

For a quantum system, described by an algebra $\mathbb{E}$ of linear operators on a Euclidean space, the q-expectations are given by a positive linear functional on $\mathbb{E}$, expressed in terms of a Hermitian density operator ρ of trace 1 as

$$\langle A\rangle = \operatorname{Tr}\rho A.$$

We refer to both the positive linear functional and the density operator as the **state** of the system. As observed in 1927 by von Neumann [216, p. 255]—who was the first to base quantum mechanics upon expectations rather than probabilities—the specification of all q-expectations uniquely determines the density operator ρ. Thus, every function of ρ can be rewritten as a function of q-expectations, and vice versa. Therefore, the notion of objective properties as functions of the state is unambiguously defined.

https://doi.org/10.1515/9783110667387-010

The state depends on how it is viewed. In this chapter, we only consider the time dependence in the Schrödinger picture,[1] so that

$$\langle A\rangle_t = \operatorname{Tr}\rho_t A. \tag{10.1}$$

A **subsystem** of a system is specified by a choice declaring some of the quantities of the system to be the distinguished quantities of the subsystem, including a choice for the Hamiltonian of the subsystem. The dynamics of the subsystem is generally not closed, hence not given by the Ehrenfest equation (2.14). Instead, one generally uses coarse-graining to obtain an approximate closed (deterministic or stochastic) dynamical description.

The state of a subsystem is completely specified by the state of any larger system it is contained in, by restricting the positive linear functional to the distinguished quantities of the subsystem. This is in contrast to the situation in traditional, state vector-based interpretations, and resolves the problems for the latter discussed in Chapter 15 of the Appendix.

In the most general form of a subsystem, no tensor product structure is assumed, unlike in traditional interpretations. However, suppose that the latter is present, so that $\mathbb{H} = \mathbb{H}_S \otimes \mathbb{H}_E$, and the quantities of the subsystem are the linear operators of $\mathbb{H}_S \otimes 1$. (This is discussed in more detail in Section 10.2.) Then, without changing any of the predictions for the subsystem, the Hilbert space of the subsystem may be taken to be the smaller Hilbert space $\mathbb{H}_S$, and the quantities of the subsystem are the linear operators of $\mathbb{H}_S$. Then the density operator of the subsystem is the **reduced state** obtained as the partial trace over the environment Hilbert space $\mathbb{H}_E$.

Only a small set of objective properties are practically (approximately) observable. Whenever we are able to compute something from raw observations according to the rules of some meaningful protocol, and it adequately agrees with something derivable from quantum physics, we call the result of that computation a **measurement** of the latter. This correctly describes the practice of measurement in its most general form.

(M) We say that a property P of a system S (encoded in its state) has been **measured** by another system, the **detector** D, if at the time of completion of the measurement and a short time thereafter (long enough that the information can be read by an observer) the detector state carries enough information about the state of the measured system S at the time when the measurement process begins to deduce with sufficient reliability the validity of property P at that time.

To give a precise formal expression for rule (M) in the context of the thermal interpretation, we have to define the property P as the validity or invalidity of a specific

1 Note that the density operator, viewed as a time-dependent object, is picture-dependent, but with the corresponding time-dependence of the linear operator A as discussed in Section 2.2, the q-expectations are picture-independent.

mathematical statement $P(\rho_S)$ about the state ρ_S of the system, and the information to be read as another specific mathematical statement $Q(\rho_D)$ about the state ρ_D of the detector. Then we have to check (theoretically or experimentally) that the dynamics of the joint system, composed of system, detector, and the relevant part of the environment implies that—with high confidence and an appropriate accuracy

$$Q(\rho_D(t)) \approx P(\rho_S(t_i)) \quad \text{for } t_f \le t \le t_f + \Delta t. \tag{10.2}$$

Here t_i and t_f denote the initial and final time of the duration of the measurement process, and Δt is the time needed to read the result.

For example, to have sufficient reasons to call the observation of a pointer position or a detector click an observation of a physical property of the measured system, one must show that (10.2) holds for the property $P(\rho_S)$ claimed to be measured and some encoding $Q(\rho_B)$ of the pointer position or detector click.

Establishing such a relation (10.2), based on experimental evidence, requires knowing already how system properties are experimentally defined, through preparation or measurement. This gives the definition of measurement the appearance of a self-referential cycle. But this self-reference is of a similar nature as the one in Callen's criterion discussed in Section 8.3, and is resolved once preparation and measurement are found to be experimentally consistent. This can be achieved in an iterative process typical for self-consistent specifications everywhere in physics.

On the other hand, deducing (10.2) theoretically is a difficult task of statistical mechanics, since the instrument is a macroscopic body that, on the fundamental level necessary for a foundation, can be treated only in terms of statistical mechanics. The investigation of this in Sections 11.6 and 11.7 will show essential differences between the traditional interpretations and the thermal interpretation.

10.2 Physical systems and their states

From a fundamental point of view, each physical system is a subsystem of the whole **universe**, the only truly isolated physical system containing the solar system. By the unitary dynamics of the universe, we have in the standard Schrödinger picture at each time t a universal density operator

$$\rho(t) = e^{-itH/\hbar}\rho(0)e^{itH/\hbar},$$

in terms of which the q-expectations (10.1) at time t are defined.

A **physical system** is a subsystem of the universe. It is selected by distinguishing the elements of a vector space $\mathbb{E}$ of quantities (linear operators on the Hilbert space of the universe) as being the quantities relevant to the subsystem, and restricting the q-expectation mapping of the universe to $\mathbb{E}$.

In many cases, the physical system S is defined by a decomposition of the Hilbert space $\mathbb{H}$ of the universe into a tensor product $\mathbb{H} = \mathbb{H}^S \otimes \mathbb{H}^E$ of a **system Hilbert space** $\mathbb{H}^S$ and an **environment Hilbert space** $\mathbb{H}^E$ for the remaining part of the universe. We call such physical systems **standard**.

Each standard physical system S has a corresponding reduced density operator, given in the standard Schrödinger picture by

$$\rho^S(t) := \mathrm{Tr}_E\, \rho(t),$$

where Tr_E denotes the partial trace over the environment. Thus, the reduced density operator $\rho^S(t)$ is the state of the physical system at time t. These are *the only* states the thermal interpretation is concerned with at all—because these are the states containing precisely the information about the q-expectations of operators of the universe attached to the system S. Indeed, the reduced density operator is defined such that for linear operators A on $\mathbb{H}^S$, describing system properties, the q-expectations are given by

$$\langle A\rangle_t := \langle A\otimes 1\rangle_t = \mathrm{Tr}\,\rho(t)(A\otimes 1) = \mathrm{Tr}\,\rho^S(t)A,$$

where 1 denotes the identity operator on $\mathbb{H}^E$. Each $\rho^S(t)$ and $\rho^S(x)$ is a Hermitian positive semidefinite linear operator on $\mathbb{H}^S$ with trace 1.

Given any Hermitian positive semidefinite linear operator ρ^S on $\mathbb{H}^S$ with trace 1, it may be possible, by utilizing the laws of Nature and the control facilities these impart on humans or machines, to ensure that at some time t_{prep} (or some spacetime position x_{prep}), $\rho^S(t_{\mathrm{prep}})$, that is, $\rho^S(x_{\mathrm{prep}})$ approximates ρ^S sufficiently well that predictions with ρ^S in place of $\rho^S(t_{\mathrm{prep}})$ or $\rho^S(x_{\mathrm{prep}})$ match experimental checks. In this case, we say that at time t_{prep} (or spacetime position x_{prep}), the system S is **prepared** in the state ρ^S. How to do this is part of the experimental art of **preparation**.

If ρ^S has rank 1, then $\rho^S = \psi\psi^*$ for some **state vector** ψ of norm one (determined by ρ^S up to a phase). In this case, we say that the system is prepared in the **pure state** ψ. Physicists can prepare a system in a pure state only when this system has very few degrees of freedom.

10.3 A single qubit as a subsystem of the universe

We consider a single qubit as a subsystem of the universe. The Hilbert space of the universe can be decomposed into a tensor product $\mathbb{H} = \mathbb{H}^S \otimes \mathbb{H}^E$ of a 2-dimensional system Hilbert space $\mathbb{H}^S$, and an environment Hilbert space $\mathbb{H}^E$ for the remaining part of the universe. We suppose that the qubit is prepared in a state defined by a general reduced density matrix ρ^S with components $\rho^S_{jk} = \langle j|\rho^S|k\rangle$. Then ρ^S is given by

$$\rho^S = \sum_{j,k} \rho^S_{jk}|j\rangle\langle k|.$$

Since ρ^S is Hermitian positive semidefinite with trace 1,

$$\rho^S = \begin{pmatrix} p & \alpha^* \\ \alpha & 1-p \end{pmatrix} \tag{10.3}$$

for some real number $p \in [0,1]$ and some complex number α with

$$|\alpha| \le \sqrt{p(1-p)}. \tag{10.4}$$

The true value of the up operator $A = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ is, according to the thermal interpretation,

$$\overline{A} = \langle A \rangle = \operatorname{Tr}_S A = p,$$

with an uncertainty of

$$\sigma_A = \sqrt{\langle A^2 \rangle - \overline{A}^2} = \sqrt{p(1-p)}.$$

In particular, the true value has no intrinsic uncertainty iff $p = 0$ or $p = 1$.

In the following, we analyze in which way p is reflected in an arbitrary environmental q-expectation. For simplicity, we assume that at preparation time $t = 0$, the density operator of the universe in the Schrödinger picture has the tensor product form

$$\rho_0 = \rho^S \otimes \rho^E = \sum_{j,k} \rho^S_{jk} |j\rangle\langle k| \otimes \rho^E. \tag{10.5}$$

(This assumption could be relaxed, but not without going through much more technical computations.) The dynamics of the universe is governed by a unitary matrix $U(t)$ turning ρ_0 into

$$\rho(t) = U(t)\rho_0 U(t)^*.$$

We may decompose $U(t)$ uniquely as

$$U(t) = \sum_{\ell,k} |\ell\rangle\langle k| \otimes U_{\ell k}(t)$$

with suitable $U_{\ell k}(t) \in \operatorname{Lin} \mathbb{H}^E$.

Let $X^E \in \operatorname{Lin} \mathbb{H}^E$ be a Hermitian quantity located in the environment, so that

$$X := 1 \otimes X^E \in \operatorname{Lin} \mathbb{H}$$

is a quantity of the universe. We want to calculate its q-expectation

$$\overline{X}_t := \langle X \rangle_t = \operatorname{Tr} \rho(t) X = \operatorname{Tr} U(t)\rho_0 U(t)^* X = \operatorname{Tr} \rho_0 U(t)^* X U(t) = \operatorname{Tr} \rho_0 X(t),$$

where

$$\begin{aligned} X(t) &= U(t)^* X U(t) = U(t)^* (1 \otimes X^E) U(t) \\ &= \sum_{\ell,j} |j\rangle\langle\ell| \otimes U_{\ell j}(t)^* (1 \otimes X^E) \sum_{\ell',k} |\ell'\rangle\langle k| \otimes U_{\ell' k}(t) \\ &= \sum_{\ell,\ell',j,k} |j\rangle\langle\ell|\ell'\rangle\langle k| \otimes U_{\ell j}(t)^* X^E U_{\ell' k}(t) \\ &= \sum_{\ell,j,k} |j\rangle\langle k| \otimes U_{\ell j}(t)^* X^E U_{\ell k}(t). \end{aligned}$$

Using (10.5), we find that

$$\begin{aligned} \overline{X}_t &= \langle X \rangle_t = \operatorname{Tr} \rho_0 X(t) \\ &= \operatorname{Tr} \sum_{j',k'} \rho^S_{j'k'} |j'\rangle\langle k'| \otimes \rho^E \sum_{\ell,j,k} |j\rangle\langle k| \otimes U_{\ell j}(t)^* X^E U_{\ell k}(t) \\ &= \sum_{\ell,j,k} \rho^S_{kj} |k\rangle\langle j| \operatorname{Tr}_E \rho^E U_{\ell j}(t)^* X^E U_{\ell k}(t). \end{aligned}$$

If we define $X^S(t) \in \operatorname{Lin} \mathbb{H}^S$ by

$$X^S(t)_{jk} := \operatorname{Tr}_E \rho^E U_{\ell j}(t)^* X^E U_{\ell k}(t), \tag{10.6}$$

we arrive at

$$\overline{X}_t = \sum_{\ell,j,k} \rho^S_{kj} |k\rangle\langle j| X^S_{jk}(t) = \operatorname{Tr}_S \rho^S X^S(t), \tag{10.7}$$

and by (10.3),

$$\overline{X}_t = p X^S_{11}(t) + (1-p) X^S_{22}(t) + 2 \operatorname{Re} \alpha X^S_{12}(t). \tag{10.8}$$

Note that in (10.7), $\overline{X}_t$ depends linearly on the system density ρ^S. In particular, when $\rho^S = \psi\psi^*$ is a pure state then $\overline{X}_t$ depends nonlinearly on ψ. As a consequence it is in general impossible to deduce from the predicted results for two initial pure states those for a superposition; additional mixed terms involving both initial pure state appear and introduce new information. Because of these mixed terms, superposition arguments break down completely – there is no superposition of measurement results.

10.4 The emergence of Born's rule

We now consider multiple preparations in the qubit state represented by ρ^S, but in multiple contexts. We label each such preparation with a label ω from some sample space Ω. Since the split into system and environment is different in each preparation,

the state ρ^E, representing the state of the environment, depends on the preparation label ω, so that $\rho^E = \rho^E(\omega)$. Since X^E was assumed to be Hermitian, (10.6) implies that the matrix $X^S(t)$ is also Hermitian and, being dependent on $\rho^E(\omega)$, depends on ω. We write

$$X^S(t) = \begin{pmatrix} \hat{x}_t(\omega) & \hat{z}_t(\omega)^* \\ \hat{z}_t(\omega) & \hat{y}_t(\omega) \end{pmatrix}$$

for the realization obtained in the preparation labeled by $\omega \in \Omega$. Then we may rewrite (10.8) as

$$\overline{X}_t = p\hat{x}_t(\omega) + (1-p)\hat{y}_t(\omega) + 2\,\mathrm{Re}\,\alpha\hat{z}_t(\omega). \tag{10.9}$$

The ω-dependence is actually a dependence on details of the environment that are uncontrollable in practice. Hence, it effectively turns $X^S(t)$ into a time-dependent random matrix and $\hat{x}_t$, $\hat{y}_t$ and $\hat{z}_t$ into time-dependent random variables, of which a new realization is obtained for each preparation of the qubit in the state represented by ρ^S. Their distribution, however, depends on more general properties of the environment and is, in principle, amenable to an analysis by the traditional techniques of statistical mechanics. Let us write

$$X^{\mathrm{eff}}(t) = \begin{pmatrix} x_t^{\mathrm{eff}} & z_t^{\mathrm{eff}*} \\ z_t^{\mathrm{eff}} & y_t^{\mathrm{eff}} \end{pmatrix}$$

for the effective mean of $X^S(t)$, averaged over all preparations $\omega \in \Omega$. As a consequence of (10.8), $\overline{X}_t$ itself behaves like a random variable, with mean

$$\overline{X}_t^{\mathrm{eff}} = \mathrm{Tr}_S\, \rho^S X^{\mathrm{eff}}(t) = \langle X^{\mathrm{eff}}(t) \rangle_S.$$

Thus, we may view every environmental q-expectation $\overline{X}_t$ as a random observation of a corresponding effective q-expectation of a quantity $X^{\mathrm{eff}}(t)$ defined on the qubit. Whereas this "randomness" is deterministic, given the complete state of the universe, it looks random due to the uncontrollable influence of the environment. Usually, $X^{\mathrm{eff}}(t)$ is just noise, and $\overline{X}_t^{\mathrm{eff}}$ is essentially zero, giving no information about the qubit. However, for specially chosen X, namely for those where X is physically related to the qubit in a significant way, $\overline{X}_t^{\mathrm{eff}}$ is nonzero and gives nontrivial statistical information about the qubit measured—it constitutes a useful measurement for a qubit property in the sense of criterion (M) from Section 10.1.

In practice, which quantity of the qubit is observed can be found out by techniques known from quantum tomography. If X depends on a parameter vector θ, then $X^{\mathrm{eff}}(t)$ also depends on θ, and we can find out the precise θ-dependence by these techniques. Therefore, we have an effective way of calibrating our measurement device. In particular, whenever we can find a value for θ, for which $X^{\mathrm{eff}}(t) = A$, we get a statistical measurement of the true value p of the up operator A.

The precise statistical properties of $X^S(t)$ can be found out by careful calibration. They can also be predicted by a theoretical analysis of the formula defining $X^S(t)$, using the standard techniques of statistical mechanics, though this may involve considerable work. Here we give an outline of how such a theoretical analysis may proceed in principle, leaving details to future investigations of particular situations amenable to a more detailed analysis.

We consider an environmental operator X^E that leads to a pointer variable $\overline{X}_t$, which moves[2] in a macroscopic time $t > 0$, a macroscopic distance to the left (in microscopic units, large negative) when $p = 0$, and to the right (large positive) when $p = 1$. In both cases, $\alpha = 0$ by (10.4). Hence, by (10.9), $\overline{X}_t = \hat{x}_t(\omega)$ in the first case, and $\overline{X}_t = \hat{y}_t(\omega)$ in the second case. Therefore,

$$\hat{x}_t(\omega) \gg 0 \gg \hat{y}_t(\omega). \tag{10.10}$$

We want to find idealized conditions, under which a measurement protocol produces measurements that follow Born's rule exactly.

In thermodynamics, we get idealized relations in the thermodynamic limit of infinite size, which are still applicable with good accuracy to systems of small but macroscopic size. Similarly, in kinetic theory, the scattering matrix, defined through an asymptotic limit of times $t \to \pm\infty$ (and the associated infinite separability of clusters) is used to define with good accuracy the collision rates and products of microscopic scattering events (where distances are small but large compared to atomic distances and times are short but large compared to the time needed to travel an atomic distance) figuring in the derivation of the kinetic equations.

This justifies in the present situation, too, that we idealize macroscopic distances and times as infinite, and therefore assume in place of (10.10) the exact but idealized limit

$$\hat{x}_t(\omega) \to \infty, \quad \hat{y}_t(\omega) \to -\infty \quad \text{for } t \to \infty. \tag{10.11}$$

The desired idealized conditions for the emergence of Born's rule are now given by the following theorem:

Theorem 10.4.1. *Let $\hat{x}_t, \hat{y}_t, \hat{z}_t$ be time-dependent random variables such that* (10.11) *holds and*

$$\hat{v}_t(\omega) := \frac{\hat{z}_t(\omega)}{\hat{x}_t(\omega)} \to 0 \quad \textit{for } t \to \infty. \tag{10.12}$$

2 This is the simplest situation amenable to a rigorous analysis. It does not suffice for the analysis of a Stern–Gerlach experiment, say, where instead one of two pointer variables at a fixed distance will show a macroscopic response. This is treated in a less rigorous fashion in Section 13.4.

If the random variable

$$u := \lim_{t\to\infty} \widehat{u}_t, \tag{10.13}$$

where

$$\widehat{u}_t(\omega) := \frac{\widehat{x}_t(\omega)}{\widehat{x}_t(\omega) - \widehat{y}_t(\omega)},$$

exists almost everywhere and is uniformly distributed in $[0,1]$ *then*

$$\Pr(\overline{X}_t \to \infty) = p, \quad \Pr(\overline{X}_t \to -\infty) = 1 - p. \tag{10.14}$$

Proof. Indeed, under the stated conditions, $\widehat{y}_t/\widehat{x}_t = 1 - 1/\widehat{u}_t$. Hence, (10.9) implies that

$$\overline{X}_t = \widehat{x}_t(p + (1-p)\widehat{y}_t/\widehat{x}_t + 2\operatorname{Re}\alpha\widehat{z}_t/\widehat{x}_t) = \widehat{x}_t(1 - (1-p)/\widehat{u}_t + 2\operatorname{Re}\alpha\widehat{v}_t).$$

Since $u \geq 0$ under our assumptions, we find in the limit that

$$\begin{aligned}\Pr(\overline{X}_t \to \infty) &= \Pr\big(1 - (1-p)/u > 0\big) = \Pr(u > 1-p) = p,\\ \Pr(\overline{X}_t \to -\infty) &= \Pr\big(1 - (1-p)/u < 0\big) = \Pr(u < 1-p) = 1-p.\end{aligned}$$

□

Therefore, in each particular experiment ω, the q-expectation $\overline{X}_t = \overline{X}_t(\omega)$ of the pointer variable will end up with a definite sign. Real detectors for microscopic events usually magnify tiny initial displacements in a single scattering event (a single escaping electron in a photomultiplier[3] or a single chemical reaction on a photographic plate) by special processes, thus making the infinite time limit irrelevant. The (here unmodeled) magnification process will lead to a unique limiting $\overline{X}$ of the same sign as $\overline{X}_t$ but macroscopically stabilized. The final measurement result $\widehat{X}(\omega)$, fully determined by $\rho_E(\omega)$, will in each case be close to one of two opposite values $\pm x$,

$$\widehat{X}(\omega) \approx \pm x. \tag{10.15}$$

Assumption (10.12) has the nature of a decoherence condition (see Section 10.5) and is likely to be satisfied under quite general conditions, using a random phase approximation argument.

Real detectors have various inefficiencies that may cause deviations from the ideal probabilistic law expressed by Born's rule. For approximately satisfying Born's rule,

3 The observable photocurrent is a property of the detector. This outcome is not quantized but a continuous burst, for each single observation of a detection event. The derivation of macroscopic electrodynamics from QED shows (cf. Section 7.4) that the measured currents are q-expectations of smeared field operators. Only the details of the smearing depend on the actual device. These are the q-expectations relevant for the thermal interpretation.

it suffices that the assumptions are satisfied only approximately. The condition that u exists and is approximately uniformly distributed in $[0,1]$ is the essential condition, which requires a thorough analysis and must be verified in each concrete setting. It is likely that in the settings treated by Allahverdyan, Balian & Nieuwenhuizen [7] and Breuer & Petruccione [54], discussed in Chapter 11, an analysis similar to that given above can be abstracted from their treatment.

10.5 Relations to decoherence

Of course, no unitary treatment of the time dependence can explain why only one of these dynamically independent components is experienced.

Erich Joos and Dieter Zeh, 1985 [150, p. 242]

According to the thermal interpretation, the state of the universe determines deterministically the outcome of each measurement, but the dependence is too sensitive to be resolved by any forseeable means. As a result, repeated experiments show a random behavior, governed under certain conditions by Born's rule.

In contrast, according to the traditional interpretations, the outcomes of experiments are identified with eigenvalues, which must be randomly selected according to Born's rule. This randomness is usually stated to be irreducible, i.e., not further analyzable. However, this is in apparent conflict with the Schrödinger equation which has no special dynamics for measurement devices. The most successful approach to relating in the traditional interpretations Born's rule to the Schrödinger equation is based on decoherence.

The fact that in many cases, the reduced density operator becomes, after an extremely short decoherence time, very close to diagonal in a suitable, environment-induced basis is called **decoherence** (Joos & Zeh [150], Zurek [318], Schlosshauer [264, 265]). In the case of measuring the up quantity of a qubit in an environment, decoherence gives as reduced density operator a diagonal matrix whose diagonal entries are the probabilities for the two outcomes up or down. The diagonalization and the resulting q-probabilities can be deduced from the core of quantum mechanics alone. However, that precisely one of the two cases actually happens cannot be deduced without using, in addition, the eigenvalue link—rule (BR5) from Section 1.1.

To see the relationship of our development with decoherence, let us review what we obtained in the previous section. In each experiment ω, the density operator $\rho_E(\omega)$ will be different, and the sign of the q-expectation (10.15) depends chaotically on the details, hence appears random.

Instead of modeling the individual case, decoherence works within the traditional interpretation, which describes everything only in statistical terms. In particular, ensembles of many identically prepared systems are considered, and the density operator ρ_E used in the decoherence approach is assumed to be an exact equilibrium

state, which is the case only in the mean over many experiments. Now the von Neumann equation governing the dynamics of the universe is linear in ρ and hence in ρ_E. The resulting reduced dynamics is in the analysis of decoherence evaluated in the Schrödinger picture and approximated by a Lindblad equation, also linear in ρ_E. By linearity, the resulting dynamics describes precisely the behavior of the mean of functions of the pointer variable in all these experiments.

Effectively, (10.9) is averaged over the ensemble, and (10.12) is replaced by the well-known effect of decoherence, the decay of the off-diagonal entries of the reduced density operator. As a result, the decoherence analysis reproduces the average results obtained from the analysis in terms of the thermal interpretation.

Thus decoherence tells roughly the same the same story as the thermal interpretation, but only in statistical terms, whereas the thermal interpretation refines this to a different, more detailed story for each single case. This is possible since in the thermal interpretation, outcomes are defined as macroscopic q-expectations approximating the microscopic quantities to be measured, and q-expectations are always single-valued. This makes a big difference in the interpretation of everything!

Note that decoherence cannot resolve the statistical picture into single events. This is because according to all traditional interpretations[4] of quantum mechanics, single events (in a single world) have no theoretical representation in the generally accepted quantum formalism. Only the thermal interpretation represents single events within the generally accepted quantum formalism, without having to assume it in the form of rule (BR5).

10.6 Measurement errors

> *Some hypotheses are dangerous, first and foremost those which are tacit and unconscious. And since we make them without knowing them, we cannot get rid of them. Here again, there is a service that mathematical physics may render us. By the precision which is its characteristic, we are compelled to formulate all the hypotheses that we would unhesitatingly make without its aid.*
>
> Henri Poincaré, 1902 [241, p. 151]

We now give a detailed analysis of the concept of measurement error. This leads to a justification and comparison of the convention used to define measurement accuracy in the thermal interpretation with the traditional convention. It is followed by an analysis of the double-slit experiment, which exemplifies the crucial differences of these conventions.

Measurement errors are ubiquitous in physical practice; their definition requires, however, some care. A single measurement produces a number, the **measurement**

4 This even holds for Bohmian mechanics. Here single events have a theoretical representation, but this representation is external to the quantum formalism, given by the additionally postulated position variables.

result. The splitting of the measurement result into the sum of an intended result (the true value) and a **measurement error** (the deviation from it) depends on what one declares to be the true value. Thus, what can be said about measurement errors depends on what one regards as the true value of something measured. This true value is a theoretical construct, an idealization arrived at by convention.

Since measurements are only actual results, never the hypothesized true values, there is no way to determine experimentally which convention is the right one. Both the quantum formalism and the experimental record are independent of what one declares to be the true value of a measurement. Different conventions only define different ways of bookkeeping, that is, different ways of splitting the same actual measurement results into a sum of true values and errors, in the communication about quantum predictions and experiments. Nothing in the bookkeeping changes the predictions and the level of their agreement with experiment.

Thus, the convention specifying what to consider as true values is entirely a matter of choice, an **interpretation**. The convention one chooses determines what one ends up with, and each interpretation must be judged in terms of its implications for convenience and accuracy. Like conventions about defining measurement units [57], interpretations can be adjusted to improvements in theoretical and experimental understanding, to better serve the scientific community.

Born's statistical interpretation of quantum mechanics gives the following convention for the prediction of measurement results for measuring a quantity given by a self-adjoint operator A: One computes a number of possible idealized measurement values, the eigenvalues of A, of which one is exactly (according to most formulations) or approximately (if level spacings are below the measurement resolution) measured, with probabilities computed from A and the density operator ρ by the probability form of Born's rule. Thus, the eigenvalues are the true values of Born's statistical interpretation.

Because of the critique of Born's rule, given in Chapter 14 of the Appendix, the thermal interpretation explicitly rejects the part of Born's rule that declares the eigenvalues of operators as the true values in a measurement. It differs from the tradition created in 1927 by Jordan, Dirac, and von Neumann, and proclaims in direct opposition the alternative convention that one computes a single, possibly idealized, measurement value, the q-expectation

$$\overline{A} := \langle A \rangle := \operatorname{Tr} \rho A$$

of A, which is approximately measured. Therefore, the true values of the thermal interpretation are the q-expectations rather than the eigenvalues.

As a result of the multivaluedness of the true values, Born's statistical interpretation needs probabilities in the very foundations of quantum physics. In contrast, as a result of the single-valuedness of the true values in the thermal interpretation, probabilities are no longer needed in the foundations of quantum physics.

Note that there is no necessity for a Gaussian error distribution. If one measure something with true value 0.37425 with a 4 digit digital device, the error distribution will be discrete, not Gaussian.

To quantitatively assess the difference in the two interpretations, we consider the **spectrum** Spec A of a linear operator on a Euclidean space $\mathbb{H}$ (a common domain of all relevant q-observables of a system), defined as the set of all $\lambda \in \mathbb{C}$ for which no linear operator $R(\lambda)$ from the completion $\overline{\mathbb{H}}$ of $\mathbb{H}$ to $\mathbb{H}$ exists, such that $(\lambda - A)R(\lambda)$ is the identity. Spec A is always a closed set.

A linear operator $A \in \operatorname{Lin}\mathbb{H}$ is called **essentially self-adjoint** if it is Hermitian and its spectrum is real (that is, a subset of $\mathbb{R}$). For N-level systems, where $\mathbb{H}$ is finite-dimensional, the spectrum coincides with the set of eigenvalues, and every Hermitian operator is essentially self-adjoint. In infinite dimensions, the spectrum contains the eigenvalues, but not every number in the spectrum must be an eigenvalue, and whether a Hermitian operator is essentially self-adjoint is a question of correct boundary conditions.

Theorem 10.6.1. *Let A be essentially self-adjoint, with value $\overline{A} := \langle A\rangle$ and q-standard deviation σ_A in a given state. Then the spectrum of A contains some real number λ with*

$$|\lambda - \overline{A}| \le \sigma_A. \tag{10.16}$$

*In particular, in the special case, where, in some state, A has a **sharp** value, defined by $\sigma_A = 0$, then the value $\langle A\rangle$ belongs to the spectrum.*

Proof. The linear operator $B = (A - \overline{A})^2 - \sigma_A^2$ is a quadratic function of A. Hence, its spectrum consists of all $\lambda' := (\lambda - \overline{A})^2 - \sigma_A^2$ with $\lambda \in \operatorname{Spec} A$; in particular, it is real. Put $\lambda_0 := \inf \operatorname{Spec} B$. Then $B - \lambda_0$ is a Hermitian operator with a real, nonnegative spectrum, hence positive semidefinite. (In infinite dimensions, this requires the use of the spectral theorem.) Thus, $B - \lambda_0 \ge 0$ and $0 \le \langle B - \lambda_0\rangle = \langle (A - \overline{A})^2\rangle - \sigma_A^2 - \lambda_0 = -\lambda_0$. Therefore, $\lambda_0 \le 0$. Since Spec B is closed, λ_0 is in the spectrum, hence has the form $(\lambda - \overline{A})^2 - \sigma_A^2$ with $\lambda \in \operatorname{Spec} A$. This λ satisfies (10.16). The final claim for $\sigma_A = 0$ follows immediately. □

Therefore, the difference between the traditional and the thermal interpretation is already within the uncertainty, and hence—from a statistical perspective—never significant. Therefore, both interpretations are in equal agreement with the experimental record. The same number obtained by a measurement may be interpreted in a dual way: It both measures some random eigenvalue to high (in the idealization even infinite) accuracy, and it simultaneously measures the q-expectation to low accuracy. In both cases, the measurement involves an additional uncertainty related to the degree of reproducibility of the measurement, given by the standard deviation of the results of repeated measurements. Tradition and the thermal interpretation agree that this

uncertainty is at least

$$\sigma_A := \sqrt{\langle A^2\rangle - \langle A\rangle^2},$$

the uncertainty appearing in Heisenberg's uncertainty relation.

10.7 What should be the true value?

A student has read such and such a number on his thermometer. He has taken no precautions. It does not matter; he has read it, and if it is only the fact which counts, this is a reality [...] Experiment only gives us a certain number of isolated points. They must be connected by a continuous line, and this is a true generalisation. But more is done. The curve thus traced will pass between and near the points observed; it will not pass through the points themselves. Thus we are not restricted to generalising our experiment, we correct it; and the physicist who would abstain from these corrections, and really content himself with experiment pure and simple, would be compelled to enunciate very extraordinary laws indeed.

Henri Poincaré, 1902 [241, pp. 142f]

As an illustration of the differences in the interpretation, we first consider some piece of digital equipment with 3-digit display, measuring some physical quantity X, using N independent measurements. Suppose the measurement results were 6.57 in 20 % of the cases, and 6.58 in 80 % of the cases. Every engineer or physicist would compute the mean $\overline{X} = 6.578$, the variance $\sigma_X^2 = \langle (X-\overline{X})^2\rangle = 0.2\cdot 0.008^2 + 0.8\cdot 0.002^2 = 16\cdot 10^{-6}$, and the standard deviation $\sigma_X = 0.004$, and would conclude that the true value of the quantity X deviates from 6.578 by an error of the order of $0.004N^{-1/2}$.

Next we consider the measurement of a Hermitian quantity $X \in \mathbb{C}^{2\times 2}$ of a 2-state quantum system in the pure up state, using N independent measurements, and suppose that we obtain exactly the same results. The thermal interpretation proceeds as before and draws the same conclusion. But Born's statistical interpretation proceeds differently and claims that there is no measurement error. Instead, each measurement result reveals one of the eigenvalues $x_1 = 6.57$ or $x_2 = 6.58$ in an unpredictable fashion with probabilities $p = 0.2$ and $1-p = 0.8$, up to statistical errors of order $O(N^{-1/2})$. For $X = \left(\begin{smallmatrix} 6.578 & 0.004 \\ 0.004 & 6.572 \end{smallmatrix}\right)$, both interpretations of the results for the 2-state quantum system are consistent with theory. However, Born's statistical interpretation deviates radically from engineering practice, without any apparent necessity.

Finally, we consider the energy measurement of an unknown system with discrete, unknown energy levels $E_1 < E_2 < \cdots$, assumed to be simple eigenvalues of the Hamiltonian. We also assume that the system is in a pure state $a_1|E_1\rangle + a_2|E_2\rangle$, where the kets denote the eigenstates of the Hamiltonian, and $|a_1|^2 = p$, $|a_2|^2 = 1-p$; for simplicity, higher levels than the lowest two are assumed to be absent. As a consequence, the q-expectation of the energy (represented by the Hamiltonian) can be exactly calculated, giving $\overline{E} = pE_1 + (1-p)E_2$. The uncertainty of the energy can be exactly calculated, too, giving $\sigma_E = \sqrt{p(1-p)}|E_1 - E_2|$.

Something analogous holds for the measurement of any quantity of an arbitrary 2-state system, such as a spin. According to the experimental record, the response of a good detector produces measurement results E concentrated at two spots of the detector, just as what one gets when measuring a classical diffusion process in a double-well potential (see, for example, Hongler & Zheng [136]). Born's rule therefore says in the present situation that the measurement results are **quantized**. This results in a bimodal distribution with two sharp peaks, with details depending on the detection method used and its resolution.

In a frequently used idealization that ignores the limited efficiency of a detector, the distribution may even be assumed to be 2-valued, with measurement results that take only one of two values E_1' and E_2', corresponding to the two modes of the bimodal distribution.

According to the thermal interpretation, each measurement result E is taken to be an approximation of the true value $\overline{E}$, with an error $|E - \overline{E}|$ of order at least σ_E. In the limit of arbitrarily many repetitions, the mean value of the approximations approaches $\overline{E}$, and their standard deviation approaches σ_E. The observed discreteness is explained as an effect due to the recording device. The latter introduces a systematic discretization error, analogous to the rounding errors in the introductory illustration. The bimodal distribution of the measurement results may be due to environment-induced randomness and environment-induced dissipation, as for a classical, environment-induced diffusion process in a double-well potential. It may also be due to the experimental setup. For example, the arrangement in a Stern–Gerlach experiment together with simple theory leads to two beams, which accounts for the approximate 2-valuedness of the response at the screen. The thermal interpretation attributes the discreteness to the detection setup, not to the true value of the spin.

According to Born's statistical interpretation in the standard formulation,[5] "the measured result will be one of the eigenvalues", each actual measurement result E is claimed to be one of the exact (in general irrational) values E_1 or E_2, and there is no measurement error.[6] However, the measurement result is not reproducible: Multiple

5 This is the formulation appearing in Wikipedia [306]. Griffiths & Schroeter [109, pp. 133] declare, "If you measure an observable [...] you are certain to get one of the eigenvalues". Peres [233, pp. 95] defines, "each one of these outcomes corresponds to one of the eigenvalues of A; that eigenvalue is then said to be the result of a measurement of A". Textbooks such as Nielsen & Chuang [218, pp. 84f] seem to avoid the issue by not referring to eigenvalues at all. But their declaration, "Quantum measurements are described by a collection $\{M_m\}$ of measurement operators. [...] The index m refers to the measurement outcomes that may occur in the experiment. [...] the probability that result m occurs", with a formula that summed over all m gives the value 1, still assumes that the values m are exact results—otherwise each of several approximations to the same intended result would have to be represented by a different M_m, and their summation would not give 1.

6 It appears unrealistic that actual measurement results can be arbitrary irrational numbers. To remedy this, one might consider a more liberal reading of the Born rule, where some additional measurement error might be acceptable. This would have the consequence that Born's rule is no longer about

repetition of the measurement results in a random sequence of values E_1 and E_2, with probabilities p and $1 - p$, respectively. In the limit of arbitrarily many repetitions, the mean value of this sequence approaches $\overline{A}$, and the standard deviation approaches σ_E.

If the energy levels are exactly known beforehand (or if the "energy" actually represents a component of a spin variable), one can calibrate the pointer scale to make $E_1' = E_1$ and $E_2' = E_2$. Then, as long as one ignores the idealization error, both interpretations become experimentally indistinguishable. However, in the more realistic case, where energy levels are only approximately known and must be inferred experimentally—the common situation in spectroscopy, the thermal interpretation, in agreement with the standard recipes for drawing inferences from inaccurate measurement results, still gives a correct account of the actual experimental situation, whereas Born's statistical interpretation paints an inadequate, idealized picture only.

measurements but about idealized measurements, whose observations are theoretical numbers, not actual results. Thus, the liberal reading of Born's rule would be a purely theoretical construct, silent about actual measurement results. I am not aware of any discussion in the published literature of such a liberal reading.

11 Measurement devices

This chapter presents the measurement problem from the point of view of the thermal interpretation of quantum physics.

Everything physicists measure is measured in a thermal environment, for which statistical thermodynamics is relevant. The thermal interpretation agrees with how one interprets measurements in thermodynamics, the macroscopic part of quantum physics, derived via statistical mechanics. By its very construction, the thermal interpretation naturally matches the classical properties of our quantum world: The thermal interpretation assigns states—and a realistic interpretation for them—to individual quantum systems, in a way that large quantum systems are naturally described by classical observables.

Since measurement devices are large quantum systems, this is the key to solving the measurement problem. We discuss the role played by macroscopic systems and the weak law of large numbers in getting readings with small uncertainty. Since quantum physics makes many deterministic predictions, for example regarding observed spectra, but also assertions about probabilities, we distinguish deterministic and statistical measurements.

To clarify the meaning of the concept of measurement, we postulate (in Section 11.1) a measurement principle that defines what it means in the thermal interpretation to measure a quantity with a specified uncertainty. This turns the measurement problem from a philosophical riddle into a scientific problem in the domain of quantum statistical mechanics, namely how the quantum dynamics correlates macroscopic readings from an instrument with properties of the state of a measured microscopic system.

The essence of scientific practice is the reproducibility of measurements, discussed in Section 11.2. Unlike in traditional interpretations, single, nonreproducible observations do not count as measurements, since this would violate the reproducibility of measurements—the essence of scientific practice. As a consequence, the measurement of a Hermitian quantity A is regarded as giving an uncertain value approximating the q-expectation $\langle A\rangle$ rather than (as tradition wanted to have it) as an exact revelation of an eigenvalue of A. This difference is most conspicuous in the interpretation of single discrete microscopic events. Except in very special circumstances, these are not reproducible. Thus, they have no scientific value in themselves and do not constitute measurement results. Scientific value is, however, in ensembles of such observations, which result in approximate measurements of q-probabilities and q-expectations.

Sections 11.3 and 11.4 distinguish deterministic and statistical measurements, depending on whether a single observation is reproducible, and discuss the role played by macroscopic systems and the weak law of large numbers in getting readings with small uncertainty. How to measure the q-probabilities of events, described in terms of

https://doi.org/10.1515/9783110667387-011

POVMs, is discussed in Section 11.5. Section 11.6 poses the measurement problem as a problem of the statistical mechanics of measurement devices, and discuss recent work by Allahverdyan, Balian, and Nieuwenhuizen on its resolution. Section 11.7 discusses the problem how, in the thermal interpretation, discrete events can arise, although q-expectations are by their very nature continuous. This is resolved by pointing to work by Breuer and Petruccione on piecewise deterministic processes (PDP). Section 11.8 sheds additional light on the origin of discrete events by an analysis of the role dissipative bistability plays in obtaining discreteness in classical physics.

11.1 Measurement protocols

The universe does not come partitioned into physical systems and measurement devices (in the following often called **instruments** or **detectors**) to measure their properties; rather these are features imposed on Nature by the scientific culture.

There does not seem to be any conceptual study that would reliably define what qualifies as a physical system or a detector. However, to draw formal conclusions about measurement in simplified model universes defined in a mathematical framework, we need at least to be able to tell precisely what constitutes a measurement *in such a model universe*. In order that these concepts get a proper meaning, they must be defined through theory in an appropriate conceptual framework.

Therefore, good foundations, including a good measurement theory, should be able to justify the informal consensus of quantum theory and its experimental practice by defining additional formal concepts about what constitutes a measurement. To be satisfactory, these must behave within the theory just as their informal relatives with the same name behave in reality. Then instrument builders may use the theory to inform themselves of what can possibly work, and instrument calibration assumes reliable laws of physics.

Thus, measurement must be grounded in theory, not—as in the traditional foundations—the other way round! In complete foundations, there are formal objects in the mathematical theory corresponding to all informal objects discussed by physicists, including those used when designing and performing measurements. Then talking about the formal objects and talking about the real objects is essentially isomorphic.

According to the thermal interpretation, properties of the system to be measured are encoded in the state of the system and its dynamics. This state and what can be deduced from it are the only objective properties of the system. On the other hand, a measuring instrument measures properties of a system of interest. The measured value—a pointer reading, a sound, a counter value, et cetera—is read off from the instrument, and hence is primarily a property of the measuring instrument and not one of the measured system. From the properties of the instrument (the instrument state), one can measure or compute the measurement results. Measurements are possible only if the microscopic laws imply quantitative relations between properties of the

measured system (that is, the system state) and the values read off from the measuring instrument (properties of the detector state).

This—typically somewhat uncertain—relation was specified in the rule (M) from Section 10.1. Considering only properties determined by a single q-expectation $\langle A \rangle$, we get as special case the principle that defines, in agreement with the general uncertainty principle (GUP) from Section 3.1 and today's NIST standard for specifying uncertainty (Taylor & Kuyatt [284]), what it means to have measured a quantity.

(MP) Measurement principle: *A macroscopic quantum device qualifies as an instrument for approximately, with uncertainty Δa, measuring a Hermitian quantity A of a system with density operator ρ, if it satisfies the following two conditions:*

(i) *(**uncertainty**) All measured results a deviate from $\overline{A}$ by approximately Δa. The measurement uncertainty is bounded below by $\Delta a \geq \sigma_A$.*

(ii) *(**reproducibility**) If the measurement can be sufficiently often repeated on systems with the same or a sufficiently similar state, then the sample mean of $(a - \overline{A})^2$ approaches Δa^2.*

Consistent with the general uncertainty principle (GUP) (from Section 3.1), the measurement principle (MP) demands that any instrument for measuring a quantity A has an uncertainty $\Delta a \geq \sigma_A$.

As customary, one writes the result of a measurement as an **uncertain number** $a \pm \Delta a$ consisting of the measured value a and its uncertainty deviation Δa, with the meaning that the error $|a - \overline{A}|$ is at most a small multiple of Δa. Because of possible systematic errors, it is generally not possible to interpret a as mean value and Δa as standard deviation. Such an interpretation is valid only if the instrument is calibrated to be unbiased.

The validity of the measurement principle for a given instrument must either be derivable from quantum models of the instrument by a theoretical analysis, or it must be checkable by experimental evidence by calibration. In general, the theoretical analysis leads to difficult problems in statistical mechanics that can be solved only approximately, and only in idealized situations. From such idealizations, one then transfers insight to make educated guesses in cases where an analysis is too difficult, and adjusts parameters in the design of the instrument by an empirical calibration process.

It is an open problem how to prove this from the statistical mechanics of measurement models. But that such a limit cannot be overcome has been checked in the early days of quantum mechanics by a number of thought experiments. Today it is still consistent with experimental capabilities and no serious proposals exist that could possibly change this situation.

The measurement principle (MP) creates the foundation of measurement theory. Physicists doing quantum physics (even those adhering to the shut-up-and-calculate mode of working) use this rule routinely, and usually without further justification. The rule applies universally. No probabilistic interpretation is needed. In particular, the first part applies also to single measurements of single systems.

Real experiments are (and must be) designed such that they allow one to determine approximately the relevant properties of the state under study, hence the values of all quantities of interest. The uncertainties in the experiments imply approximations, which, if treated probabilistically, need an *additional* probabilistic layer. Indeed, the formulation "at least of the order of σ_A" allows for the frequent situation that the measurement uncertainty is larger than the intrinsic (theoretical) uncertainty σ_A. Expectations from this secondary layer, which involve probabilistic statements about situations that are uncertain due to neglected, but in principle observable, details (see Peres [232]), happen to have the same formal properties as the values on the primary layer, though their physical origin and meaning is completely different. Thus, **probability** and **statistics** have the same role as in classical physics, namely as the *art and science of interpreting collections of data obtained by measurements,* rather than figuring in the foundations.

According to (MP), exact measurements have $\Delta a = 0$, and hence $\sigma_A = 0$. This indeed happens in the measuring of systems in a pure state when the state vector is an eigenstate of the quantity measured. Therefore, part of Born's rule holds: *Whenever a quantity A is measured exactly,*[1] *its value is an eigenvalue of A*. But for inexact (that is, almost all) measurements, the thermal interpretation rejects Born's rule as an axiom, defining what counts as a measurement result. With this move, all criticism from Chapter 14 of the Appendix becomes void, since Born's rule remains valid only in a limited validity; see Section 11.8.

11.2 Statistical and deterministic measurements

To understand the precise meaning of the notion of measurement, we look at measurements in the context of classical physics and chemistry. There are two basic kinds of measurements, destructive measurements and nondestructive measurements.

Nondestructive measurements either leave the state of the object measured unchanged (such as in the measurement of the length of a macroscopic object), or modify it temporarily during the measurement (for example, temporarily deforming it to measure the stiffness) in such a way that the object returns to its original state after the measurement is completed. **Destructive measurements** permanently change the state of the object measured, usually by destroying all or part of it during the measurement process. Examples are the determination of the age of an archeological artifact by dendrochronology, or many traditional methods of finding the chemical composition of a material.

1 Note that the discrete particle spin measurements in a Stern–Gerlach experiment, say, are not exact measurements in the present sense. They measure exactly only a coarse-grained position on the screen, but are very low-accuracy measurements of the particle spin q-expectations; see Section 13.4.

In both cases, the measurement gives some **posterior** information about the **prior** state, that is, the state of the object before the start of the measurement process. In the case of destructive measurements, it usually also gives some information about the products of the destruction, from which properties of the prior states are deduced by reasoning.

A characteristic context of destructive measurements is the presence of a large, sufficiently homogeneous object. Tiny parts of it are subjected to destructive measurements to discover their relevant properties. The homogeneity of the object then implies that the properties deduced from the destructive measurements are also properties of the remainder of the object. Thus, destructive measurements of a tiny fraction of a homogeneous object give information about the whole object, including its unmeasured part. By our definition, we obtain in this way a nearly nondestructive measurement of the whole object.

Alternatively, a large number of essentially identical objects are present, a few of which are subjected to a destructive measurement. The results of the measurement are then taken as being representative of the properties of the unmeasured objects. In case the measurements on the objects measured do not agree, one can still make statistical statements about the unmeasured objects, approximately valid within the realm of validity of the law of large numbers. However, this no longer gives valid information about a single unmeasured object, but only information about the whole population of unmeasured objects. Thus, one may regard the measurement on multiple trial objects as a measurement of the state of the whole population.

Similar to classical destructive measurements, experiments measuring individual particles change (unless specially tuned to be "non-demolition measurements") the state of the individual particles in an unpredictable way. Just as in a destructive measurement, their precise state before measurement can never be ascertained. Only probabilities for their collective behavior can be given by averaging over many observations of different realizations.

For example, the analysis of experimental particle collisions is based on measuring the momentum and charge of many individual collision products. But *individual* collision events (the momentum and charge of the individual collision products) are not predicted by the theory—only the possibilities and their collective statistics, their distribution in a *collection* of equally prepared events. Indeed, probabilities mean nothing for a single collision. What does it mean that the particular collision event recorded at a particular time in a particular place is obtained with probability 0.07? Nothing objective—the single collision simply happened and has no associated probability. A scientific statement about probabilities is always a statement about a process that can be repeated many times under essentially identical conditions.

The requirement (MP) for a measuring instrument includes the reproducibility of the resulting measurement values. Reproducibility in the general sense that all systems prepared in the same state have to behave alike when measured is a basic requirement for all natural sciences. The term "alike" has two different interpretations

depending on the context: Either "alike" is meant in the deterministic sense of "approximately equal within the specified accuracy". Or "alike" is meant in the statistical sense of "approximately reproducing in the long run the same probabilities and mean values". An object deserves the name "instrument" only if it behaves in one or the other of these ways.

Corresponding to the two meanings, we distinguish two kinds of measuring instruments, deterministic ones and statistical ones. Consequently, the quantitative relationship between the system state and the measurement results may be deterministic or statistical, depending on what is measured.

Radioactive decay, when modeled on the level of individual particles, is a typical **statistical** phenomenon. It needs a stochastic description as a branching process, similar to classical birth and death processes in biological population dynamics. The same holds for particle scattering, the measurement of cross sections, since particles may be created or annihilated, and for detection events, such as recording photons by a photoelectric device or particle tracks in a bubble chamber.

On the other hand, although quantum physics generally counts as an intrinsically probabilistic theory, it is important to realize that it not only makes assertions about probabilities, but also makes many **deterministic** predictions verifiable by experiment. These deterministic predictions fall into two classes:

(i) Predictions of numerical values believed to have a precise value in Nature:
   - The most impressive proof of the correctness of quantum field theory in microphysics is the magnetic moment of the electron, predicted by quantum electrodynamics (QED) to the phenomenal accuracy of 12 significant digit agreement with the experimental value. It is a universal constant, determined solely by the two parameters in QED, the electron mass and the fine structure constant.
   - QED also predicts correctly emission and absorption spectra of atoms and molecules, both the spectral positions and the corresponding line widths.
   - Quantum hadrodynamics allows the prediction of the masses of all isotopes of the chemical elements in terms of models with only a limited number of parameters.

(ii) Predictions of qualitative properties, or of numerical values believed to be not exactly determined, but which are accurate with a tiny, computable uncertainty.
   - The modern form of quantum mechanics was discovered through its successful description and organization of a multitude of spectroscopic details on the position and width of spectral lines in atomic and molecular spectra.
   - QED predicts correctly the color of gold, the liquidity of mercury at room temperature, and the hardness of diamond.
   - Quantum physics enables the computation of thermodynamic equations of state for a huge number of materials. Equations of state are used in engineering in a deterministic manner, with negligible uncertainty. Engineers usually

need not explicitly consider quantum effects, since these are encoded in their empirical formulas for the equations of states.
- Quantum chemistry predicts correctly rates of chemical reactions.
- From quantum physics, one may also compute transport coefficients for deterministic kinetic equations used in a variety of applications.

Thus, quantum physics makes both deterministic and statistical assertions, depending on which system it is applied to and on the state or the variables to be determined. Statistical mechanics is mainly concerned with deterministic prediction of class (ii) in the above classification.

Predictions of class (i) are partly related to spectral properties of the Hamiltonian of a quantum system, and partly to properties deduced from form factors, which are deterministic byproducts of scattering calculations. In both cases, classical measurements account adequately for the experimental record.

The traditional interpretations of quantum mechanics do only rudimentarily address the deterministic aspects of quantum mechanics, requiring very idealized assumptions (being in an eigenstate of the quantity measured) that are questionable in all deterministic situations described above.

11.3 Macroscopic systems and deterministic instruments

A **macroscopic system** is a system large enough to be described sufficiently well by the methods of statistical mechanics,[2] where, due to the law of large numbers, one obtains essentially deterministic results.

The weak law of large numbers implies that quantities averaged over a large population of identically prepared systems become highly significant when their value is nonzero, even when no single quantity is significant. This explains the success of Boltzmann's statistical mechanics to provide an effectively deterministic description of ideal gases, where all particles may be assumed to be independent and identically prepared.

In real, nonideal gases, the independence assumption is only approximately valid because of possible interactions, and in liquids, the independence is completely lost. The power of the statistical mechanics of Gibbs lies in the fact that it allows replacing simple statistical reasoning on populations, based on independence by more sophisticated algebraic techniques that give answers even in extremely complex interacting cases. Typically, the uncertainty is of the order $O(N^{-1/2})$, where N is the mean number

2 However, as discussed by Sklar [275], both the frequentist and the subjective interpretation of probability in statistical mechanics have significant foundational problems, already in the framework of classical physics. These problems are absent in the thermal interpretation, where single systems are described by mixed states, without any implied statistical connotation.

of identical microsystems making up the macroscopic system. Thus, the thermal interpretation associates with macroscopic objects, essentially classical quantities, whose uncertain value (q-expectation) has a tiny uncertainty only.

In particular, the macroscopic pointer of a measurement instrument always has a well-defined position, given by the q-expectation of the Heisenberg operator $x(t)$ corresponding to the center of mass of its $N \gg 1$ particles at time t. The uncertain pointer position at time t is $\langle x(t)\rangle \pm \sigma_{x(t)}$, where the q-expectation is taken in the Heisenberg state of the universe (or any sufficiently isolated piece of it). Therefore, the position is fully determined by the state of the pointer—but it is an uncertain position. By the law of large numbers, the uncertainty $\sigma_{x(t)}$ is of order $N^{-1/2}$. Typically, this limit accuracy is much better than the accuracy of the actual reading. Thus, we get well-defined pointer readings, leading within the reading accuracy to deterministic measurement results.

Whether by this or by other means, whenever one obtains an essentially deterministic measurement result, we may say that measuring is done by a deterministic instrument, defined as follows:

A **deterministic instrument** is a measuring instrument that measures objective properties, deterministic functions of the state ρ of the system measured, within some known margin of accuracy, in terms of some property read from the instrument, a macroscopic system. A special case is the measurement of a quantity A, since the uncertain value $\overline{A} = \operatorname{Tr}\rho A$ of A is a function of the state ρ of the system. Thus, if measurements yield values $a \approx \overline{A}$ within some uncertainty Δa, the corresponding instrument is a deterministic instrument for **measuring** A within this accuracy.

Not all reliably measurable macroscopic properties are naturally given as q-expectations. For example, in equilibrium thermodynamics, temperature T, pressure P, and chemical potential μ have no simple description in terms of microscopic variables. They figure only as a parameter in the expression for the grand canonical phase space density $\rho = e^{-(H+PV-\mu N)/kT}$ of the state. But T and P are computable from ρ via the thermodynamic formalism of statistical mechanics, and hence are objective properties in the sense of the thermal interpretation. The definition (M), from Section 10.1, of what it means to measure something therefore still applies. More generally, it applies (see Neumaier & Westra [214]) to arbitrary macroscopic thermal systems in equilibrium, whose state is characterized by a collection of finitely many extensive and intensive thermodynamic variables related by the standard thermodynamic relations, expressed in terms of an equation of state for the materials making up the thermal system.

In particular, the measurement of temperature and pressure of, say, a single brick of iron in equilibrium is a perfectly sensible special case of our definition (M) of what it means to measure something. On the other hand, according to the traditional interpretations, they are not even "observables"—although they are observable in any meaningful sense of the word!

11.4 Statistical instruments

The measurement of a tiny, **microscopic** system, often consisting of only a single particle, is of a completely different nature. Now the uncertainties do not benefit from the law of large numbers, and the relevant quantities often are no longer significant, in the sense that their uncertain value is already of the order of their uncertainties. In this case, the necessary quantitative relations between properties of the measured system and the values read off from the measuring instrument are only visible as stochastic correlations.

The results of single measurements are no longer reproducibly observable numbers. In the thermal interpretation, a single detection event is therefore not regarded as a measurement of a property of a measured microscopic system, but only as a property of the macroscopic detector correlated to the nature of the incident fields.

This is the essential part, where the thermal interpretation differs from tradition. Indeed, from a single detection event, one can only extract very little information about the state of a microscopic system. Conversely, from the state of a microscopic system, one can usually predict only probabilities for single detection events.

All readings from a photographic image or from the scale of a measuring instrument, done by an observer, are deterministic measurements of an instrument property by the observer. Indeed, what is measured by the eye is the particle density of blackened silver on a photographic plate, or that of iron of the tip of the pointer on the scale, and these are extensive variables in a continuum mechanical local equilibrium description of the instrument.

The historically unquestioned interpretation of such detection events as the measurement of a particle position is one of the reasons for the failure of traditional interpretations to give a satisfying solution of the measurement problem. The thermal interpretation is here more careful and treats detection events instead as a statistical measurement of particle beam intensity.

To obtain comprehensive information about the state of a single microscopic system is therefore impossible. To collect enough information about the prepared state, and hence the state of a system measured, one needs either time-resolved measurements on a single stationary system (available, for example, for atoms in ion traps or for electrons in quantum dots), or a population of identically prepared systems. In the latter case, one can get useful microscopic state information through quantum tomography; see Section 11.5.

Thus, in case of measurements on microscopic quantum systems, the quantitative relationship between measurement results and measured properties only takes the form of a statistical correlation. The reproducibly observable items, and hence the carrier of scientific information, are statistical mean values and probabilities. These are indeed predictable by quantum physics. But—in contrast to the conventional terminology applied to single detection events for photons or electrons—*the individual events no longer count as definite measurements of single system properties*.

This characteristics of the thermal interpretation is an essential difference to traditional interpretations, for which each event is a definite measurement.

A **statistical instrument** determines its final measurement results from a large number of raw measurements by averaging or by more advanced statistical procedures, often involving computer processing. Again, due to the law of large numbers, one obtains essentially deterministic results, but now from very noisy raw measurements. Examples include low-intensity photodetection, the estimation of probabilities for classical or quantum stochastic processes, astronomical instruments for measuring the properties of galaxies, or the measurement of population dynamics in biology.

This behavior guarantees reproducibility. In other words, systems prepared in the same state behave in the same way under measurement—in a deterministic sense for a deterministic instrument, and in a statistical sense for a statistical one. In both cases, the final measurement results approximate, with limited accuracy, the value of a function F of the state of the system under consideration.

11.5 Probability measurements

> *When it uses probabilities, [...] science regards them [...] as measurable (and calculable) physical quantities like lengths, energies, and wavelengths. [...] The probability of a truly single event is intrinsically unmeasurable and [...] science has nothing to say about [it.] To obtain the value of a physical quantity, one must measure it a number of times. Each measurement contains an error, and the 'true' value is (usually) computed as the arithmetic mean of all measured values. [...] in a similar way [we] measure the relative frequency [...] of an event in a series of trials. Each relative frequency contains an error, and the 'true' probability is computed as the mean of the relative frequencies over a number of series. [...] nothing strange or inconsistent is left in the idea of probability as a measurable physical quantity.*
>
> Henry Margenau, 1950 [179, pp. 250–252]

By its definition, the notions of q-expectations and q-probabilities belong to the formal core of quantum mechanics and are independent of any interpretation. But in the thermal interpretation all q-expectations, and in particular all q-probabilities, are among the objective properties. We discuss here basic aspects of their measurement.

By the law of large numbers, q-expectations can be measured with (in principle) arbitrarily high accuracy by taking sample means of low-accuracy measurements, whenever there is a device (the **preparation**)—for example, a particle accelerator—that produces a large number of independent copies (realizations) of the same quantum system. The accuracy improves by a factor of $\sqrt{N}$, where N is the sample size.

In the same way, the q-probabilities p are approximately measurable as relative frequencies. As a consequence of the weak law of large numbers (3.11), the uncertainty of the **relative frequency** p_N, defined as the sample mean of ideal binary measurements in a sample of N independent realizations of the statement P, is $\sigma = \sigma_P\sqrt{N} = \sqrt{p(1-p)/N}$. This uncertainty approaches zero when the sample size N gets arbitrarily

large. Thus, measuring a probability by a relative frequency gives (in principle) arbitrarily accurate results.

This gives a fully adequate operational definition of probabilities without any logical problems of the same quality as operational definitions of highly accurate length or time measurements. The q-probabilities are theoretical observables; they are measured as relative frequencies, to some reasonable accuracy that can be quantified by the associated uncertainty. We draw conclusions about sufficiently uncertain situations based on observed sample means and relative frequencies on a sample of significant size, and we quantify our remaining uncertainty by statistical safeguards (confidence intervals, et cetera), well knowing that these sometimes fail. For example, the 5σ**-rule** for the discovery of elementary particles[3] tries to guard against such failures. This view of probabilities as measurable entities is the one described in the above quote by Margenau (if 'final' is read for 'true').

Thus, probability has an objective interpretation precisely to the extent that objective protocols for taking the sample measurements, that is, how to distinguish a positive from a negative test, are agreed upon. Any subjectivity remaining lies in the question of deciding which protocol should be used for accepting a measurement as 'correct'. Different protocols may give different results. Both classically and quantum mechanically, the experimental context needed to define the protocol influences the outcome. In particular, there is a big difference between the description of an event before it occurs (*predicting* it) or after it occurs (*analyzing* it). This is captured rigorously in classical probability theory by conditional probabilities (see Section 3.4), and less rigorously in quantum physics by the so-called **collapse** of the wave function, where the description of the state of a particle is different before and after it passes a filter (polarizer, magnet, double slit, et cetera). Thus, we may view the collapse as the quantum analogue of the change of conditional probability when the context changes due to new information.

Measurements in the form of discrete events (such as the appearance of clicks, flashes, or particle tracks) may be described in terms of an **event-based instrument** characterized by a discrete family of possible measurement results $a_1, a_2, \dots$ that may be real or complex numbers, vectors, or fields, and nonnegative Hermitan quantities $P_1, P_2, \dots$ satisfying

$$P_1 + P_2 + \cdots = 1. \tag{11.1}$$

The nonnegativity of the P_k implies that all q-probabilities

$$p_k = \langle P_k \rangle = \operatorname{Tr} \rho P_k \tag{11.2}$$

3 The 5σ-rule requires for a new particle evidence showing a deviation of at least 5σ from the best results predicted with an uncertainty of σ without the new particle.

are nonnegative, and (11.1) guarantees that the q-probabilities always add up to 1. By its definition, the notion of q-probabilities belongs to the formal core of quantum mechanics and is independent of any interpretation.

Unlike in all traditional interpretations, the thermal interpretation considers the observable result a_k not as the exact measurement result of some "observable" with counterintuitive quantum properties, but as a (due to the tiny sample size very low accuracy) statistical measurement of some q-expectation.

In the thermal interpretation all q-expectations, and in particular, all q-probabilities are among the objective properties. As described in Section 11.5, a q-probability p may be approximately measured as relative frequency, whenever there is an event-generating device (the **preparation**) that produces a large number N of independent copies (realizations) of the same quantum system. In this case, we require that if the measured system is in the state ρ, the instrument gives the observable result a_k with a relative frequency approaching the q-probability p_k as the sample size gets arbitrarily large.

An **event-based instrument** is a statistical instrument measuring the probability of events modeled by a discrete (classical or quantum) statistical process. In the quantum case, it is mathematically described by a **positive operator-valued measure**, short **POVM**, defined as a family $P_1, P_2, \dots$ of Hermitian, positive semidefinite operators satisfying (11.1) (or a continuous generalization of this).

POVMs originated around 1975 in work by Helstrom [123] on quantum detection and estimation theory and are discussed in some detail in Peres [232]. They describe the most general quantum measurement of interest in quantum information theory. Which operators P_k correctly describe a statistical instrument can, in principle, be found out by suitable **calibration measurements**. Indeed, if we feed the instrument with enough systems prepared in known states ρ_j, we can measure approximate probabilities $p_{jk} \approx \langle P_k \rangle_j = \operatorname{Tr} \rho_j P_k$. By choosing the states diverse enough, one may approximately reconstruct P_k from this information by a process called **quantum tomography**. In quantum information theory, the Hilbert spaces are finite-dimensional. Hence, the quantities form the algebra $\mathbb{E} = \mathbb{C}^{N\times N}$ of complex $N \times N$ matrices. In this case, the density operator is the **density matrix** ρ, a complex Hermitian $N \times N$-matrix with trace one, together with the trace formula

$$\langle A \rangle = \operatorname{Tr} \rho A.$$

Since $\langle 1 \rangle = 1$, a set of $N^2 - 1$ binary tests for specific states, repeated often enough, suffices for the state determination. Indeed, it is easy to see that repeated tests for the states e^j, the unit vectors with just one entry 1 and other entries 0, tests the diagonal elements of the density matrix, and since the trace is 1, one of these diagonal elements can be computed from the knowledge of all others. Tests for $e^j + e^k$ and $e^j + ie^k$ for all $j < k$ then allow the determination of the (j, k) and (k, j) entries. Therefore, frequent repetition of a total of $N - 1 + 2\binom{N}{2} = N^2 - 1$ particular tests determines the full state.

The optimal reconstruction to a given accuracy, using a minimal number of individual measurements, is the subject of **quantum estimation theory**, still an active frontier of research.

Distinguished from a stochastic instrument performing event-based measurements is an **event-based filter**, which turns an input state ρ with probability

$$p_k := \langle R_k^* R_k \rangle$$

into an output state

$$\rho_k := \frac{1}{p_k} R_k \rho R_k^*.$$

Here the R_k are operators satisfying

$$\sum_k R_k^* R_k = 1.$$

Each case possible may be considered as a potential event; the collection of possible events is then described by the POVM with $P_k := R_k^* R_k$.

11.6 Chaos, randomness, and quantum measurement

The coarse-graining involved in reducing the unitary dynamics of the universe to the dissipative dynamics of an open system, described by a limited collection of relevant quantities, leads to the neglect of many high-frequency details. This results in stochastic features, either in the models themselves, or in the relation between models and Nature. In this way, the deterministic Ehrenfest dynamics of all q-expectations of the universe gives rise to stochastic features at the coarse-grained level. The quantitative derivation of the stochastic properties is therefore reduced to a problem of quantum statistical mechanics.

Deterministic coarse-grained models are usually chaotic, introducing a second source of randomness on a more tangible level. To explain the randomness inherent in the measurement of quantum observables in a qualitative way, it seems to be sufficient to invoke the chaoticity of the coarse-grained approximations to equations of motion of a system, including both the measured system and its measuring device.

Many coarse-grained models are chaotic. In particular, we have seen in Section 6.1 how the coherent action principle (the Dirac–Frenkel variational procedure applied to coherent states) gives coarse-grained approximations that only track a number of relevant variables and exhibit chaotic behavior. A more well-known coarse-grained example are the Navier–Stokes equations, used in practice to model realistic fluid flow. They are well known to be chaotic and though deterministic in principle, they exhibit in practice stochastic features that make up the phenomenon of turbulence. In general, deterministic chaos, as present in classical mechanics, results in empirical randomness.

According to the thermal interpretation, this chaotic effective motion in reduced description is responsible for the probabilistic features of quantum mechanics. The empirical randomness is taken to be an emergent feature of deterministic chaos implicit in the deterministic Ehrenfest dynamics of the universe discussed in Section 2.2. Since the Ehrenfest dynamics is linear, it seems to be strange to consider it chaotic. However, the chaotic nature appears once one restricts attention to the macroscopically relevant q-expectations, where the influence of the ignored q-expectations is felt as a stochastic contribution to the effective coarse-grained dynamics of the relevant q-expectations.

The dynamics we actually observe is the quantum dynamics of a more complex system, coarse-grained to a dynamics of these few degrees of freedom—at increasing level of coarse-graining described by Kadanoff–Baym equations, Boltzmann-type kinetic equations, and hydrodynamic equations, such as the Navier–Stokes equations. These coarse-grained systems generally behave like classical dynamical systems with regimes of highly chaotic motion.

In general, deterministic chaos manifests itself once one uses a coarse-grained, locally finite-dimensional parameterization of the quantum states. This leads to an approximation where, except in exactly solvable systems, the parameters characterizing the state of the universe (or a selected part of it) change dynamically in a chaotic fashion.

As discussed in more detail in Section 6.8, Zhang & Feng [315] used the coherent action principle, restricted to group coherent states, to get a coarse-grained system of ordinary differential equations approximating the dynamics of the q-expectations of macroscopic operators of certain multiparticle quantum systems. At high resolution, this deterministic dynamics is highly chaotic. Whereas this study makes quite special assumptions, it illustrates how although the basic dynamics in quantum physics is linear, chaotic motion results once attention is restricted to a tractable approximation. This chaoticity is indeed a general feature of coarse-graining approximation schemes for the dynamics of q-expectations or the associated reduced density functions. (For a discussion of quantum chaos from a completely different perspective, see Peres [232, p. 353ff] and the survey by Haake [111].)

According to the thermal interpretation, quantum physics is the basic framework for the description of objective reality (including everything reproducible studied in experimental physics), from the smallest to the largest scales. In particular, quantum physics must give an account of whatever happens in an experiment, when both the equipment and the systems under study are modeled on the quantum level. In experiments probing the foundations of quantum physics, one customarily observes a small number of field and correlation degrees of freedom (often simplified in a few particle setting) by means of macroscopic equipment. To model the observation of such a tiny quantum system by a macroscopic detector, one must simply extend the coarse-grained description of the detector by adding a few additional quantum degrees of freedom for the measured system, together with the appropriate interactions.

The metastability needed for a reliable quantum detector (for example, in a bubble chamber) together with chaoticity then naturally leads to a random behavior of the individual detection events.

In terms of the thermal interpretation, the **measurement problem**—how to show that an experimentally assumed relation between measured system and detector results is actually consistent with the quantum dynamics—becomes a precise problem in quantum statistical mechanics.[4] Of course, details must be derived in a mathematical manner from the theoretical assumptions inherent in the formal core.

A number of recent papers by Allahverdyan, Balian & Nieuwenhuizen (in the following short **AB&N**), reviewed in Neumaier [201], addressed this issue. Here we only discuss AB&N's paper [7]. This paper carefully analyzes the assumptions regarding the statistical mechanics used that actually go into the analysis in their long, detailed paper [6]. The latter discusses an only slightly idealized, but on the whole realistic measurement process, formulated completely in terms of quantum dynamics.

To avoid circularity in their arguments, AB&N introduce the name **q-expectation value** for $\langle A\rangle := \operatorname{Tr}\rho A$ considered as a formal construct rather than a statistical entity, and similarly (as we do in Footnote 2 from Section 3.1) q-variance and other q-notions, to be distinguished from their classical statistical meaning. This allows them to use the formalism of statistical mechanics without any reference to prior statistical notions. The statistical implications are instead derived from the analysis within this formal framework (together with explicitly specified interpretation rules), resulting in a derivation of Born's rule and the time scales, in which the implied correlations of microscopic states and measurement results are dynamically realized, based on a unitary dynamics of the full quantum system involving the microscopic system, the measurement device, and a heat bath modeling the environment.

Most important for the interpretation in [7] is AB&N's "interpretative principle 1":

ABN principle: If the q-variance of a macroscopic observable is negligible in relative size, its q-expectation value is identified with the value of the corresponding macroscopic physical variable, even for an individual system.

This is just a special case of the basic uncertainty principle central to the thermal interpretation of quantum physics!

11.7 The statistical mechanics of definite, discrete events

Generally in physics, invariance and the resulting reproducibility determine what counts as an objective property. In 3-dimensional vision, observed length is not a property of an observed object by itself, but a property of the object and the distance

4 When one insists on the rigid, far too idealized framework in which quantum physics was developed historically, and in which it is typically introduced in textbooks, the measurement problem is instead an ill-posed, vexing philosophical riddle.

from and the orientation relative to the observer. Extrapolation to zero distance viewed orthogonally defines an invariant objective length. In relativity, length is a property of the object, the distance from and the orientation relative to the observer. and in addition the relative speed to the observer. Again, suitable extrapolation defines the invariant objective length.

Science is about reproducible aspects of our world, and hence not all permanent records, but only reproducible results count as measurement results. This is the main difference between the thermal interpretation and traditional interpretations of quantum mechanics.

As a consequence, a measurement of a Hermitian quantity A gives an uncertain value approximating the q-expectation $\langle A\rangle$ rather than (as tradition wanted to have it) an exact eigenvalue of A. This difference is most conspicuous in the interpretation of single discrete events. Since most single microscopic observations are not reproducible, they have no scientific value in themselves, and do not constitute measurement results.[5] Scientific value is, however, in ensembles of such observations, which result in approximate measurements of q-probabilities and q-expectations.

In the thermal interpretation, the traditional difficulty to show that there is always a unique outcome is trivially solved, since by definition, the outcome of reading a macroscopic quantity is its expectation value, with negligible uncertainty. Instead, we now have a new difficulty absent in traditional interpretations: An explanation is required why, although fed with a stationary interaction, certain detectors record random individual events!

For example, why does a low-intensity beam of light produce in a photodetector a discrete signal? The uncertain observed value is the q-expectation of a photocurrent, which a priori has a continuum of possible values. But observed are two clearly different regimes that allow one to clearly distinguish between the occurrence and the nonoccurence of a detection event. In the thermal interpretation, we do not consider the single detector event as a property of the observed beam ("a particle arrived through the beam"), since only the statistics of an ensemble of detector events (for example, a Poisson distribution of the number of events in some large time interval) is reproducible, and hence constitutes an objective property of the beam. But why these discrete events can be clearly distinguished at all needs an explanation.

Section 6.6 of the book on open quantum systems by Breuer & Petruccione [54, pp. 348–350] (in the following short **B&P**) addresses this issue. The dynamics of a large quantum system, consisting of an observed system and a detector observing it, is treated there as a classical dynamical system for the density operator with stochastic initial conditions, and reduced by appropriate coarse-graining to a classical stochas-

5 The same holds in classical stochastic models. If die casting is part of a stochastic system description, the single die cast tells nothing about the state of the model, and hence is of no value for the scientific study of the model.

tic equation for the coarse-grained stochastic density operator. The derivation is done using standard assumptions from classical statistical mechanics only, in the same way as one would proceed in statistical mechanics for any other classical dynamical system.

The detector must include enough of the environment to produce irreversible results (and hence determines what is read out). B&P model the latter by assuming separated time scales and the validity of the Markov approximation. Both assumptions hold only when the detector is big enough to be dissipative. (The latter is typically achieved by including in the detector a heat bath consisting of an infinite number of harmonic oscillators.) Since B&P make these assumptions without deriving them, their analysis holds for general dissipative detectors. But—as always in statistical mechanics—one must check for any concrete application that these assumptions are plausible.

In sufficiently idealized settings, these assumptions can actually be proved rigorously, but this is beyond the scope of the treatment by B&P. Rigorous results (without the discussion of selective measurement, but probably sufficient to establish the assumptions used by B&P) were first derived by Davies 1974 and later papers with the same title. See also the detailed survey by Spohn [276].

The stochastic equations discussed by B&P preserve the rank of the density operator, and hence can be applied to pure states, where the dynamics reduces in general to that of a piecewise deterministic stochastic process (PDP), a diffusion process, or a combination of both. The piecewise deterministic part accounts for the statistics of **discrete events**.

In the cases treated by B&P in their Chapter 6 (usually for the pure case only), the PDP corresponds to photodetection, which measures the particle number operator (with a discrete spectrum); the diffusion processes correspond to homodyne or heterodyne detection, which measure quadratures (with a continuous spectrum). B&P obtain the latter from the PDP by a limiting process in the spirit of the traditional approach treating a continuous spectrum as a limit of a discrete spectrum.

But although the pointer reading is a position measurement of the pointer, what is measured about the particle is not its position, but the variable correlated with the pointer reading—the photon number or the quadrature. Particle position is as indeterminate as before. Indeed, investigation of the PDP process shows that the collapsed states created by the PDP are approximate eigenstates of the number operator or the quadrature. Thus, the PDP can be interpreted in Copenhagen terms as constituting the repeated measurement of particle number or quadrature.

For photodetection, one gets at the end a PDP for the reduced state vector, only using classical probabilities in the whole derivation. But after everything has been done, the PDP may be interpreted in terms of quantum jumps, without having postulated any irreducible "collapse" as in the Copenhagen interpretation (see Section 16.2 of the Appendix). This suggests that, in general, collapse in a single observed system—in the modern POVM version of the von Neumann postulates for quantum dynamics—is

derivable from the unitary dynamics of a bigger system under the standard assumptions that go into the traditional derivations in classical statistical mechanics.

The arguments show that to go from unitarity to irreversible discrete events in Hamiltonian quantum mechanics, one does not need to assume more than to go from reversibility to irreversibility in Hamiltonian classical mechanics—namely a suitable form of the Markov approximation. Statistical assumptions are not needed to make pointers acquire a well-defined position or to create photocurrents; the standard dissipation arguments are enough. This gives stochastic equations for definite macroscopic outcomes.

11.8 Dissipation, bistability, and Born's rule

The development of B&P is mathematical and quantitative, but abstract. But there is also a qualitative explanation why the discreteness that makes its appearance in quantum mechanics actually quite natural, explained by environment-induced randomness and the associated environment-induced dissipation. This provides a more intuitive view of how the thermal interpretation settles this foundational key issue. (For generalities about environment-induced randomness and dissipation see, for example, the first two chapters of Calzetta & Hu [59].)

In general, dissipation in the effective, human time scales dynamics of a set of relevant variables is a frequent situation even when the fully detailed dynamics is conservative. This effective dissipation is the reason underlying the possibility of reduced, coarse-grained descriptions whenever there is a separation of time scales for slow and fast processes. Then one can coarse-grain by eliminating the fast modes and obtain a simpler limiting (effectively time-averaged) description on the slow manifold, the manifold where all slow motion happens (see, for example, Lorenz [176], Roberts [248]). Whenever the slow manifold is disconnected, metastable states of the full manifold decay under uncontrollable (environment-induced) perturbations into states in one of the connected components of the slow manifold. The components thus label random events selected by environmental noise.

For example, consider bending a classical, rotationally symmetric rod using a force in the direction of the axis of the rod. If the force exceeds the threshold, where the straight rod becomes metastable only, the rod will bend into a random, but definite direction. The randomness arises from the classical Hamiltonian dynamics together with the tiniest amount of noise causing a deviation from perfect symmetry. The same analysis can be made for the dynamics of a metastable inverted classical pendulum.

Similarly, perturbing in an uncontrollable way a classical bistable system arbitrarily little from the intermediate metastable state linking the two local minima of the potential leads to a tiny random move into one of the two potential wells. Even the slightest amount of dissipated energy fixes the selection of the potential well, and

more dissipation forces the system after a short relaxation time to be very close to one of the two minimum position. This is the principle underlying the emergence of chemical reactions of molecules (recognizable bound states of few atoms) from a multiparticle atomic description in transition state theory (HÄNGGI et al. [112]).

Papers on optical bistability (for example, DRUMMOND & WALLS [74], STEYN-ROSS & GARDINER [278]) show how coarse-grained bistability arises from a quantum model by projecting out irrelevant degrees of freedom. Any bistable system obtained as a reduced description from a larger unitary system behaves in the same way. Thus, one expects a few-particle quantum system, coupled to a macroscopic metastable instrument, to behave in the same way when (as is usual) unstable stationary points are present.

Therefore, within the accuracy of the approximations involved, bistability and more general multistability, leads together with dissipation to the emergence of random discrete events from deterministic dynamics. The timescale of the emergence of these discrete events is likely to be a small multiple of the decoherence time of the system; see SCHLOSSHAUER [265].

The traditional introductory textbook approach to measurement is based on the concept of ideal measurements—illustrated with Stern–Gerlach experiments, low-density double-slit experiments, and the like. These experiments primarily illustrate an antiquated view of measurement, dating back to the time before 1975, when POVMs (see Section 11.5) were still unknown. Until then, quantum measurements used to be described solely in terms of ideal statistical measurements. These constitute a special case (or for continuum measurements a special limiting case) of POVMs, where the P_k form a family of **orthogonal projectors**, that is, linear operators satisfying

$$P_k^2 = P_k = P_k^*, \quad P_j P_k = 0 \quad \text{for } j \neq k,$$

to the eigenspaces of a self-adjoint quantity A (or the components of a vector A of commuting such quantities) with discrete spectrum given by $a_1, a_2, \ldots$. We may call a statistical instrument for measuring A in terms of such a POVM a **Born instrument**, and the instrument is then said to perform an **ideal measurement** of A.

Ideal measurements of A have quite strong theoretical properties since under the stated assumptions, the instrument-based statistical average

$$\overline{f(A)} = p_1 f(a_1) + p_2 f(a_2) + \cdots$$

agrees for all functions f, defined on the spectrum of A with the model-based value $\langle f(A) \rangle$. In an ideal measurement, the relationship between the properties of the instrument and the properties of the system have a purely correlative nature, and the rule (11.2) defining the probabilities reduces to the discrete form (for example, DRUMMOND & WALLS [74], STEYN-ROSS & GARDINER [278]) of **Born's rule**. On the other hand, these strong properties are bought at the price of idealization, since (unlike more general POVMs) they frequently result in effects incompatible with real measurements.

As we saw above, bistability explains the appearance of discrete binary events. Once these are given, they provide an ideal binary measurement of the statement associated with the event—giving on a single event the result 0 or 1 with a large uncertainty compared with the probability it measures. The weak law of large numbers then implies that the relative frequencies in sufficiently large samples approximate the probability for a positive event, here given by Born's rule for ideal binary measurements.

In particular, for $P_1 = \phi\phi^*$, where ϕ has norm 1, and $P_2 = 1 - P$, this covers the case whether a quantum system in the pure state ψ responds to a test for state ϕ, and gives Born's squared probability amplitude formula $p = |\phi^*\psi|^2$ for the probability of a positive test result. When interpreted as a measure of beam intensity, this formula is identical with Malus' law from 1809 (see Section 8.6).

12 Particles

> *Before Maxwell people thought of physical reality – in so far as it represented events in nature – as material points, whose changes consist only in motions which are subject to total differential equations. After Maxwell they thought of physical reality as represented by continuous fields, not mechanically explicable, which are subject to partial differential equations. This change in the conception of reality is the most profound and the most fruitful that physics has experienced since Newton; but it must also be granted that the complete realisation of the programme implied in this idea has not by any means been carried out yet. [...] The last and most successful creation of theoretical physics, quantum mechanics, differs fundamentally in its principles from the two programmes which we will briefly designate as Newton's and Maxwell's. For the quantities which appear in its laws lay no claim to describe physical reality itself but only the probabilities for the occurrence of one of the physical realities to which attention is being directed. [...] I am inclined to think that physicists will not be satisfied in the long run with this kind of indirect description of reality, even if an adaptation of the theory to the demand of general relativity can be achieved in a satisfactory way.*
>
> Albert Einstein, 1931 [78]

As we have seen in Section 8.4, the basis of our perception of an objective reality is statistical mechanics and field theory. These are routinely used in equilibrium and nonequilibrium statistical thermodynamics (see, for example, Calzetta & Hu [59]), which govern the macroscopic laws of everyday life. Thus the quantum theory of fields must be the fundamental description of Nature. Therefore, the simpler quantum mechanics of particles is necessarily a derived description.

This chapter discusses the extent to which a particle picture of matter and radiation is appropriate. In relativistic quantum field theory, particles appear only as asymptotic excitations of the fields. The quantum particle concept therefore makes mathematical (and intuitive) sense only under very special circumstances, namely in those where a system actually behaves like particles do:

- As quasiparticles, they are useful in the domain, where the geometric optics perspective applies.
- Before and after a scattering event in a dilute gas or a particle collider ring, particles can be considered as being essentially free. During the very short interaction time itself, the scattering process cannot be described by a particle picture, but needs a hydromechanical description or a kinetic description via Kadanoff–Baym equations.

Already at the very beginning of modern quantum mechanics, when Born found his probability interpretation, the scattering process of single particles in an external classical field was understood (see Section 14.1 of the Appendix) to be a wave process relating initially and finally stationary particles, freely moving in their center of mass frame.

https://doi.org/10.1515/9783110667387-012

It is a historical accident that one continues to use the name particle in the many microscopic situations, where it is grossly inappropriate to think of it with the classical meaning of a tiny bullet moving through space. When we restrict the use of the particle concept to where it is appropriate, or when we do not think of particles as "objects", in both cases all mystery is gone, and the foundations become fully rational. The inappropriate focus on the particle aspect of quantum mechanics created the appearance of mystery; common sense is restored by focussing instead on the field aspect.

To substantiate the above summary view, we give in Section 12.1 an analysis of the photoelectric effect and in Section 12.2 a review of the origin of particle tracks. This leads in Section 12.3 to a discussion about the reality of particles. In the remainder of the chapter, we discuss the relation between fields and particles in more detail. Section 12.4 poses the problem how the nonrelativistic multiparticle description arises from relativistic quantum field theory. Section 12.5 considers the case of free quantum fields, and Section 12.6 the case of interacting quantum fields. Section 12.7, the final section, discusses the semiclassical view of quasiparticles in the approximation familiar from geometric optics.

12.1 The photoelectric effect

> *For many purposes the quantization of the electromagnetic field is not necessary at all, and the response of the photodetector can be understood even if we continue to picture the field in terms of classical electromagnetic waves, provided the photoelectrons are treated by quantum mechanics. The field then simply behaves as an external potential that perturbs the bound electrons of the photocathode. Such an approach to the problem is sometimes known as semiclassical [...] Of course, it has certain limitations, and if pushed too far the semiclassical treatment will reveal some internal contradictions. However, that does not detract from its usefulness in many circumstances. As we shall see, for those electromagnetic fields for which an adequate classical description exists, the semiclassical and the fully quantized treatments of the photodetection problem yield virtually identical answers.*
>
> Leonard Mandel and Emil Wolf, 1995 [178, p. 439]

In quantum optics experiments, both sources and beams are extended macroscopic objects describable by quantum field theory and statistical mechanics, and hence have (according to the thermal interpretation) associated nearly classical observables—densities, intensities, correlation functions—computable from quantum physics in terms of q-expectations.

An instructive example is the **photoelectric effect**, the measurement of a classical free electromagnetic field by means of a photomultiplier. Faint coherent laser light falling on a photosensitive plate causes randomly placed detection events following a Poisson distribution. Conventionally, this effect is ascribed to the particle nature of light, and each detection event is taken as a proof that a photon arrived. Upon closer

analysis, however, the detection events are found to be artifacts caused by the quantum nature of the photosensitive plate. This must be the case because the analysis can be done in a model of the process, in which no photons exist. Such an analysis was first done in 1926 by Wentzel [298] and is discussed in modern textbooks, for example, in Sections 9.1–9.5 of Mandel & Wolf [178], a standard reference for quantum optics.

Classical input to a quantum system is conventionally represented in the Hamiltonian of the quantum system by an interaction term containing the classical source as an external field or potential. In the semiclassical analysis of the photoelectric effect, the detector is modeled as a many-electron quantum system, whereas the incident light triggering the detector is modeled as an external electromagnetic field. The result of the analysis is that if the classical field consists of electromagnetic waves (light), with a frequency exceeding some threshold, then the detector emits a random stream of photoelectrons with a rate that, for not too strong light, is proportional to the intensity of the incident light. The predictions are quantitatively correct for normal light. The response of the detector to the light is statistical, and only the rate (a short time mean) with which the electrons are emitted bears a quantitative relation with the intensity. Thus, the emitted photoelectrons form a statistical measurement of the intensity of the incident light.

In the model, the electron field of the detector responds to a classical external electromagnetic radiation field by emitting electrons according to Poisson-law probabilities. Therefore, the quantum detector produces discrete Poisson-distributed clicks, although the source is completely continuous. The state space of this quantum system consists of multielectron states only. Thus, the multielectron system (followed by a macroscopic decoherence process that leads to the multiple localization of the emitted electron field) is responsible for the creation of the discrete detection pattern.[1]

The results of this analysis are somewhat surprising: Although the semiclassical model used to derive the quantitatively correct predictions does not involve photons at all, the discrete nature of the electron emissions implies that a photodetector responds to classical light as if it were composed of randomly arriving photons! (The latter was the basis for the original explanation of the photoeffect, for which Einstein received the Nobel prize.)

Though only an approximation to the quantum electromagnetic field, the classical external field discussed so far shows that the discrete response of a photodetector cannot be due to its interactions with particles, or more generally not to the quantum nature of the detected object. The discrete response is due to the detector itself,

1 This was already clearly expressed in 1924 by Jeans [147], who writes on pp. 80: "The fundamental law of quantum-dynamics, that radiant energy is emitted and absorbed only in complete quanta, is no longer interpreted as meaning that the ether can carry radiant energy only in complete quanta, but that matter can deliver or absorb radiant energy only by complete quanta."

and triggered by the interaction with a field. A field mediating the interaction must be present with sufficient intensity to transmit the energy necessary for the detection events.

Both a classical and a quantum field produce such a response. Quantum electrodynamics is of course needed to explain special quantum effects of light revealed in modern experiments, but not for the photoelectric effect. Indeed, finer analysis reveals that beams in nonclassical states may give a counting statistics significantly different from that of the classical analysis. But this only shows that the beam description needs quantum field theory (where people conventionally use the language of photons), not that there must be actual particles called photons. Only the quantitative details change in the case of quantum fields, but nothing depends on the presence or absence of "photons".

The present analysis constitutes a proof that detection events happen in the detector without photons being present. Hence, one cannot tell from a detection event whether the cause was a photon or a classical field. But if the detectors cannot even distinguish in the theoretical analysis an external classical field from an impinging photon, based on which analysis should an experimentor decide?

The only sensible conclusion is that photons are figurative properties of quantum fields manifesting themselves only in the detectors. Before detection, there are no photons; one just has beams of light in an entangled state. The beams are far more real than the photons that they are supposed to contain. This is consistent with Stokes' classical view of the qubit discussed in Section 8.6.

This shows the importance of differentiating between prepared states of the system (here beams of classical or quantum light) and measured events in the instrument (here the amplified emitted electrons). The measurement results are primarily a property of the instrument, and their interpretation as a property of the system measured needs theoretical analysis to be conclusive.

12.2 Particle tracks

In the theory of radioactive disintegration, as presented by Gamow, the α-particle is represented by a spherical wave which slowly leaks out of the nucleus. On the other hand, the α-particle, once emerged, has particle-like properties, the most striking being the ray tracks that it forms in a Wilson cloud chamber. It is a little difficult to picture how it is that an outgoing spherical wave can produce a straight track [...] no mention should be made of the α-ray being a particle at all.

Nevill Mott, 1929 [187, p. 79]

The number of cracks produced by a projectile hitting a glass sheet provides information on the impactor's speed and the properties of the sheet. [...] A projectile traveling at 22.2 meters per second generates four cracks in a 1-millimeter-thick sheet of Plexiglas. [...] A 56.7 meter-per-second projectile generates eight radial cracks in the same thickness Plexiglass sheet as above.

Michael Schirber, 2013 [262]

It may seem that the reality of individual *massive* particles is established beyond doubt through the observation of particle tracks in bubble chambers and other path-tracking devices. But are the observed "tracks" guaranteed to be traces of particles?

The paper by Schirber [262] discusses essentially the same phenomenon in a fully classical context, where a bullet is fired into a sheet of glass and produces a large number of radial cracks in random directions, shown in the first figure there. (See also Falcao & Parisio [84] and Vandenberghe & Villermaux [288].)

In this case, the discrete, random detection events (the cracks) are a manifestation of broken symmetry when something impacts a material that—unlike water—cannot respond in a radially symmetric way. Randomness is inevitable in the breaking of a radial symmetry into discrete events. The projectile creates an outgoing spherical stress wave in the plexiglas and produces straight cracks. In fact, once initiated, the growth of a crack in a solid is not very different from the growth of a track in a bubble chamber, except that the energies and time scales are quite different. Only the initiation is random.

Would observed tracks in a high-energy collision experiment prove without doubt the existence of particles, one would have to conclude that the projectile contains "crack particles", whose number is a function of the energy of the projectile—just as the number of photons in a laser beam is a function of its energy, and the number of events produced by a laser beam hitting a photodetector provides information on the impactor's brightness. Only the details are different.

Therefore the number of discrete detection events cannot be regarded as obvious evidence for the existence of the same number of associated invisible objects. They are at best evidence of the impact of something.

It is strange that in the classical experiment with the bullet, we see broken symmetry due to microscopic uncertainty, whereas quantum tradition claims that in a bubble chamber we see irreducible quantum randomness.

How do we know whether the tracks in a bubble chamber do not have a similar origin as the classical cracks? In both cases, something impinging on the detector produces a collection of traces. Though the details are different, the fundamental mechanism appears to be the same. In both cases, there is a complicated macroscopic process that breaks the symmetry and produces tracklike events. Thus, there is no a priori reason why in one case, but not the other, the lines should be interpreted as evidence of particles.

Tracks in a bubble chamber are also a manifestation of broken symmetry when a radially symmetric α particle field produced by a radioactive nucleus impacts a bubble chamber. A famous paper by Mott [187] (see also Figari & Teta [87, 66]) explains in detail how in a bubble chamber complete particle tracks appear in random directions, because of the discrete quantum nature of the bubble chamber—nowhere is made use of the particle nature of the impacting radial wave! Thus, whereas the details are quantum mechanical, the underlying principle is classical!

What we see in a bubble chamber are droplets condensing due to ionization caused by a local piece of a spherical wave emanating from a radioactive nucleus. Mott analyzes the impact of the spherical wave and proceeds without reference to anything outside the quantum formalism. He shows (pp. 80) that, in the absence of a deflecting magnetic field, the atoms cannot both be ionized, unless they lie in a nearly straight line with the radioactive nucleus. Mott needs Born's rule only for interpreting the final outcome in terms of probabilities and finds it consistent with a distribution of straight path only. This fully explains the tracks, without making any claims about position measurements or particle pointer states or collapse assumptions. There is no direct reference to the α particle causing the ionizations.

Mott's analysis suggests that after the collision, the scattered part forms a spherical wave (and not particles flying in different directions) until the wave reaches the detector. The spherical wave is nowhere replaced by flying particles. This makes his analysis very close to a field theoretical treatment. Particles appear to be ghostlike, and only macroscopic (hence field-like) things are observed.

12.3 How real are particles?

> *Von solchem Licht aber genügt ein einziges Lichtquant, um das Elektron völlig aus seiner 'Bahn' zu werfen (weshalb von einer solchen Bahn immer nur ein einziger Raumpunkt definiert werden kann), das Wort 'Bahn' hat hier also keinen vernünftigen Sinn. [...] Die 'Bahn' entsteht erst dadurch, daß wir sie beobachten."*
>
> Werner Heisenberg, 1927 [117, p. 176, p. 185]

> *[...] regard the state operator ρ as the fundamental description of the state generated by the thermal emission process, which yields a population of systems each of which is a single electron.*
>
> Leslie Ballentine, 1998 [22, p. 240]

Everything said in the previous two sections supports the view that, except during the detection event, particles are unreal in the sense of having no associated objective properties, and that they have a shade of reality only as identical realizations of a population.

This conclusion that particles are unreal and have meaning only during measurement or figuratively as part of a population is also reflected in the statistical interpretation of quantum mechanics championed by Ballentine [23, Chapter 9], who denies that a single system has a state, and instead assigns the state to a population of similarly prepared systems. By assigning objective properties to the preparation procedure, to the beam, but not to the electrons, he effectively declares beams, but not electrons, as real.

Thus, in Ballentine's view, what really exists are the beams and the discrete effects they produce when subjected to measurement. But expressed in this way, this is an

intuitive picture completely orthogonal to the traditional interpretations of quantum mechanics.

Our conclusion is also reflected in the Copenhagen interpretation of quantum mechanics, which says that an unmeasured photon has no path, hence (in terms of objective properties) is unreal. This view that here the word 'path' has no definable meaning, that the 'path' comes into being only when we observe it, was introduced by Heisenberg [117].

In quantum gravity, the particle concept even becomes observer-dependent, due to the Unruh effect.

Even before the advent of quantum mechanics, it turned out that, in classical statistical mechanics, atoms are indistinguishable not only due to practical limitations, but in principle, and that there is no theoretically conceivable way to distinguish them as individuals—if it were possible, the resulting predictions would have an additional entropy of mixing, which is in conflict with the observed thermodynamical properties of bulk systems. This means that there are fundamental constraints that forbid the atoms in a classical multiparticle system to have individual properties. Thus, in a classical multiparticle system, the atoms are anonymous objects without an identity, and as discussed in Section 3.2, expectations—averages over all indistinguishable particles of each kind—are the only classical observables. This situation persists in the quantum case, where atoms and elementary particles are in principle[2] indistinguishable, too.

Thus, the atoms and elementary particles in a multiparticle quantum system are also anonymous objects without an identity. This is reflected in the fact that on the physical Hilbert space of correctly symmetrized wave functions, no particle position operator is definable; particle positions are spurious objects. The definable operators are cumulative N-particle operators. When expressed in terms of the second quantization formalism, these become linear combinations of quantum field operators and their products. On the level of fields, indistinguishability is obvious: A field with two indistinguishable local excitations near x and y is the same as a field with the two local excitations near y and x. Thus, the q-expectations of quantum field operators are the natural generalizations of the classical observables.

Finally, the black body spectrum, whose determination by Planck [238] began the era of quantum physics and was the motivation for Einstein [77] to develop the notion of a light particle. However, the black body spectrum was explained in 1924 by Bose [48] through the canonical ensemble of what is now called a Bose–Einstein gas. For the equilibrium thermodynamics of a Bose-Einstein gas one today only need the

2 Exceptions are cases, where the range of some quantity identifies a unique particle. (This is analogous to the identifiability of outliers in anonymous statistical data.) Examples include a single atom prepared in an ion trap, single atoms on the surface of some other material, the atom closest to a given lattice position in a piece of metal, or, in the Hartree-Fock approximation, the outermost electron of an atom. In this case, the identification can be made by expectations of quantities containing a characteristic function of the defining property as a factor.

maximum entropy state corresponding to the q-expectation of a Hamiltonian of the uncoupled harmonic oscillators of a free field, with their discrete eigenvalues. In the derivation, only macroscopic q-expectations are involved; no reference to particles is needed.

We may conclude from our discussion that the lack of reality of quantum particles is well supported by the literature, in many different ways. Turning the suggestions obtained into definite statements of the thermal interpretation, we may summarize our findings as follows:

- Quantum fields (beams, et cetera) are real and have associated objective properties, given by q-expectation values.
- Quantum particles (photons, α particles, et cetera) do not exist, except when measured. They are detection events created by the detector and mediated by fields.
- On the other hand, a semiclassical picture of particles often works from a phenomenological point of view. How this comes about in an approximate way is the theme of the remainder of this chapter.

12.4 Particles from quantum fields

If QFT is about fields, how can its restriction to nonrelativistic phenomena be about particles?
Art Hobson, 2013 [134, p. 212]

In physics practice, it is often unavoidable to switch between representations featuring different levels of detail. The fundamental theory of elementary particles and fields, with the most detailed description, is quantum field theory. Since quantum field theory is fundamental, the simpler quantum mechanics of particles is necessarily a derived description.

How to obtain the quantum mechanics of particles from relativistic interacting quantum field theory is a nontrivial problem. The traditional textbook description in terms of scattering and associated propagators does not give a description at finite times. On the other hand, books on molecular quantum electrodynamics such as Liu [175] or Salam [257], which need to work with a finite time dynamics, make heuristic approximations and do not really derive the particle picture from QED.

In the fundamental reality, represented by objective properties of quantum field theory, expressed at finite times in hydrodynamic terms, fields concentrated in fairly narrow regions move along uncertain flow lines determined by effective field equations.

In the particle description, these fields are somehow replaced by a quantum mechanical model of moving particles. The uncertainty is now accounted for by the uncertain value of the position $\mathbf{q}(t)$ of each particle together with its uncertainty $\sigma_{\mathbf{q}(t)}$, at any time t, providing not a continuous trajectory, but a fuzzy world tube defining their location. The momentum of the quantum particles is also uncertain. For example, the momentum vector of a particle at CERN is measured by collecting information

from many responding wires and applying curve fitting techniques to get an approximate curve of positions at all times, and inferring from its derivative an uncertain momentum. Similar techniques are used for particle tracks on photographic plates or in bubble chambers.

How one finds from a relativistic quantum field description of a beam a corresponding quantum mechanical particle description has hardly received attention so far. Whereas informally, particles are considered to be elementary excitations of the quantum fields, this can be given an exact meaning only for free field theories. In interacting relativistic quantum fields, the notion is, at finite times, approximate only.

That the approximation problem is nontrivial can be seen from the fact that in quantum field theory, position is a certain parameter, whereas in the quantum mechanics of particles, position is an uncertain quantity. Thus, in the approximation process, position loses its parameter status and becomes uncertain. How, precisely, is unknown.

12.5 Fock space and particle description

A precise correspondence between particles and fields is possible only in free quantum field theories. These are described by distribution-valued operators on a Fock space. The latter is completely determined by its 1-particle sector, the single particle space.

Poincaré invariance, locality, and the uniqueness of the vacuum state imply that the single particle space of a free quantum field theory furnishes a causal unitary irreducible representation of the Poincaré group. These representations were classified in 1939 by Wigner [304]. This is why particle theorists say that elementary particles are causal unitary irreducible representations of the Poincaré group. Thus, elementary particles are something exceedingly abstract, not tiny, fuzzy quantum balls!

For spin ≤ 1, these representations happen to roughly match the solution space of certain wave equations for a single relativistic particle in the conventional sense of quantum mechanics, but only if one discards the contributions of all negative energy states of the latter. In relativistic quantum field theory, the latter reappear as states for antiparticles—a different kind of particles with different properties. This already shows that there is something very unnatural about the relativistic particle picture on the quantum-mechanical single-particle level.

In general, a field description on the particle level in terms of a conventional multiparticle structure is necessarily based on a Fock space representation with a number operator N with spectrum consisting precisely of the nonnegative integers. The eigenspace for the eigenvalue 1 of N then defines the bare single-particle Hilbert space. In the relativistic case, the resulting description is one in terms of bare, unphysical particles.

Untangling the S-matrix using bare perturbation theory replaces the real-time dynamics of the quantum fields by a nontemporal infinite sum of contributions of multivariate integrals, depicted in shorthand by Feynman diagrams showing a web of virtual particles. The Feynman diagrams provide a pictorial representation of the formalism of bare perturbation theory. Free real particles show as external lines, whereas the interaction is represented in terms of internal lines, figuratively called virtual particles. Most of the resulting integrals (all except the tree diagrams) are infinite and physically meaningless. A renormalization process turns the sum of all diagrams with a fixed number of loops (where the infinities cancel) into finite numbers, whose sum over not too high orders (the series is asymptotic only) has an (approximate) physical meaning. But in the renormalization process the intuitive connection of the lines depicted in Feynman diagrams—the alleged world lines of virtual particles, in the popular myth (see Neumaier [199])—gets completely lost. Nothing resembles anything like a process in time; described by the theory and the computations is only a black box probabilistic model of the in-out behavior of multiparticle scattering.

12.6 Physical particles in interacting field theories

> *All our knowledge concerning the internal properties of atoms is derived from experiments on their radiation or collision reactions, such that the interpretation of experimental facts ultimately depends on the abstractions of radiation in free space, and free material particles. [...]*
> *The use of observations concerning the behaviour of particles in the atom rests on the possibility of neglecting, during the process of observation, the interaction between the particles, thus regarding them as free. [...]*
> *The wave mechanical solutions can be visualised only in so far as they can be described with the aid of the concept of free particles. [...]*
> *Summarising, it might be said that the concepts of stationary states and individual transition processes within their proper field of application possess just as much or as little 'reality' as the very idea of individual particles.*
>
> Niels Bohr, 1927 [39, pp. 586–589]

> *In its mature form, the idea of quantum field theory is that quantum fields are the basic ingredients of the universe, and particles are just bundles of energy and momentum of the fields.*
>
> Steven Weinberg, 1997 [296, p. 2]

Whereas the conventional construction of relativistic quantum field theories starts with Fock space, a relativistic interacting quantum field itself cannot be described in terms of a Fock space. The Fock space structure (and hence the particle structure) of the initial scaffolding is destroyed by the necessary renormalization, since the number operator cannot be renormalized. Only the asymptotic fields figuring in the S-matrix reside in a Fock space—for colored quarks, because of confinement, not even in a conventional Fock space with a positive definite inner product, but only in an indefinite Fock–Krein space.

As a consequence, the particle concept is only asymptotically valid, under conditions where particles are essentially free. Traditionally, the discussion of particle issues in relativistic interacting quantum fields is therefore restricted to scattering processes involving asymptotical particle states. Only the S-matrix provides meaning to quantum particles, in an asymptotic sense, describing Born's rule for scattering processes. In the formulation of Section 14.1 of the Appendix: In a scattering experiment described by the S-matrix S

$$\Pr(\psi_{\text{out}}|\psi_{\text{in}}) := |\psi_{\text{out}}^* S \psi_{\text{in}}|^2$$

is the conditional probability density that scattering of particles prepared in the in-state ψ_{in} results in particles in the out-state ψ_{out}.

Indeed, textbook scattering theory for elementary particles is the *only* place where Born's rule is used in quantum field theory. Here the in- and out-states are asymptotic eigenstates of total momentum, labelled by a maximal collection of independent quantum numbers (including particle momenta and spins). An **asymptotic quantity** is a q-observable still visible in the limits of time $t \to \infty$ or $t \to -\infty$, so that scattering theory says something interesting about it. This is relevant since quantum dynamics is very fast, but measurements take time. Measuring times are already very well approximated by infinity, on the time scale of typical quantum processes. Thus, only asymptotic quantities have a reasonably well-defined response. That is why information about microsystems is always collected via scattering experiments described by the S-matrix, which connects the asymptotic preparation at time $t = -\infty$ and the asymptotic measurement at time $t = +\infty$. Particle momenta (like other conserved additive quantities) are asymptotic quantities.

In quantum field theory, scattering theory is just the special case of a universe containing only a tiny number of particles with known momentum at time $t = -\infty$, whose behavior at time $t = +\infty$ is to be predicted. This caricature of a universe is justified only when the few-particle system is reasonably well isolated from the remainder of the universe. In a real experiment, this is a good approximation to a collision experiment when the length and time scale of a collision is tiny compared to the length and timescale of the surrounding preparation and detection process. Much care is taken in modern colliders to achieve this to the required degree of accuracy.

12.7 Semiclassical approximation and geometric optics

I still believe in the possibility of giving a model of reality, a theory, that is to say, which shall represent events themselves and not merely the probability of their occurrence. On the other hand, it seems to me certain that we have to give up the notion of an absolute localization of the particles in a theoretical model. This seems to me to be the correct theoretical interpretation of Heisenberg's indeterminacy relation. [...] Only if this sort of representation of the atomistic structure be obtained could I regard the quantum problem within the framework of a continuum theory as solved.

Albert Einstein, 1934 [79, p. 169]

In the preceding, we discussed the precise notion of particles in relativistic quantum field theory—an asymptotic notion only. Cross-sections for the scattering processes computed in this way are supposed to be exact (assuming the idealization that the underlying theory is exact and the computations are done exactly).

However, the particle picture has another very practical use, as an approximate, semiclassical concept valid whenever the fields are concentrated along a single (possibly bent) ray and the resolution is coarse enough. When these conditions apply, one is no longer in the full quantum domain and can already describe everything semiclassically, that is, classical with small quantum corrections. Thus, the particle concept is useful when and only when the semiclassical description is already adequate. Whenever one uses the particle picture beyond scattering theory (and in particular always when one has to interpret what people using the particle language say), one silently acknowledges that one works in a semiclassical picture, where a particle description makes approximate sense, except during collisions.

A **particle** is a blop of high-field concentrations well-localized in phase space (that is, in the kinetic approximation of quantum field theory), with a boundary, whose width (or the width in transversal directions for a moving particle) is tiny compared to its diameter.

Thus, field concentrations must be such that their (smeared) density peaks at reasonably well-defined locations in phase space. At this point, similar to the regime of geometric optics for classical electromagnetic fields, these peaks behave like particles. Therefore, particles are approximately defined as local excitations of a field, and they have (as wavelets in classical mechanics) an uncertain (not exactly definable) position. Their (necessarily approximate) position and momentum behaves approximately classically (and gives rise to a classical picture of quantum particles) in the regime corresponding to geometric optics. When the spatial resolution is such that the conditions for the applicability of geometric optics hold, particles can be used as an adequate approximate concept.

In a collision experiment, it is valid to say that particles travel on incoming and outgoing beams in spacetime, whereas they are far apart, since this is a good semiclassical description of the free particles in a paraxial approximation. But when they come close, the semiclassical description breaks down and one needs full quantum field theory to describe what happens.

The exact state of the interacting system is now a complicated state in a renormalized quantum field Hilbert space[3] that no one so far was able to characterize; it is only known (Haag's theorem) that it cannot be the asymptotic Fock space describing the noninteracting particles. Since it is not a Fock space, talking about particles during the interaction makes no longer sense—the quantum fields of which the particles are

3 This Hilbert space is generally nonseparable, a direct sum of the uncountably many Hilbert spaces corresponding to the different superselection sectors.

elementary excitations become very nonparticle-like. After the collision products separate well enough, the semiclassical description becomes feasible again, and one can talk again about particles traveling along beams.

Thus, whereas the field picture is always valid, the picture of particles traveling along beams or other world tubes is appropriate except close to the collision of two world tubes. The behavior there is effectively described in a black box fashion by the S-matrix. This is a reasonable approximation if the collision speed is high enough, so that one can take the in- and outgoing particles as being at time $-\infty$ and $+\infty$, and can ignore what happens at finite times, that is, during the encounter. Thus, in the **semiclassical** description, we have between collisions real particles described by asymptotic states, whereas the collisions themselves—where the particle picture no longer makes sense—are described using a black box view featuring the S-matrix. To calculate the S-matrix, one may work in renormalized perturbation theory, using quantum field theory.

Using the intuition of geometric optics requires a locally free effective description. In a locally homogeneous background, such an effective description is usually achievable through the introduction of **quasiparticles**. These are collective field modes that propagate as if they were free. If the composition of the background changes, the definition of the quasiparticles changes as well.

In particular, the photons in glass or air are quasiparticles conceptually different from those in vacuum. Similarly, the moving electrons in a metal are quasiparticles conceptually different from those in vacuum. This shows that photons, electrons, and other elementary particles have no conceptual identity across interfaces. A photon, traditionally taken to be emitted by a source, then passing a system of lenses, prisms, half-silvered mirrors, and other optical equipment, changes its identity each time it changes its environment!

This is corroborated by the field of **electron optics**, where geometric rays are used to calculate properties of magnetic and electrostatic lenses for electron beams.

Problems abound if one tries to push the analogies beyond the semiclassical domain of validity of the particle concept. Already in classical relativistic mechanics, point trajectories are idealizations, restricted to a treatment of the motion of a single point in a classical external field. By a result of Currie et al. [63], classical relativistic multiparticle point trajectories are inconsistent with a Hamiltonian dynamics. Thus, one should not expect them to exist in quantum physics either. They are appropriate only as an approximate description.

Note that this semiclassical domain of validity of the particle picture excludes experiments with multilocal fields generated by beam-splitters, half-silvered mirrors, double slits, diffraction, long-distance entanglement, and the like. It is there where the attempt to stick to the particle picture leads to all sorts of counterintuitive features. But these are caused by the now inadequate particle imagery, not by strange features of quantum field theory itself.

13 Some quantum experiments

In this chapter, we look in more detail at some of the traditional quantum experiments from the point of view of the thermal interpretation.

An analysis of individual events on a screen leads in Section 13.1 to the picture of a quantum bucket for measuring a continuous variable with a device capable only of producing discrete results. The notion of quantum currents from Section 7.4 may be used to visualize in the thermal interpretation the finite time dynamics of particle decay (Section 13.2) and the Stern–Gerlach experiment (Section 13.3). Another analysis of the Stern–Gerlach experiment, now in terms of particles, is given in Section 13.4. The discrepancy between the thermal interpretation and traditional interpretations is seen to be of order $O(\hbar)$. Finally, in the context of entanglement experiments, we discuss notions of causality (Section 13.5) and nonlocality (Section 13.6) and their relation to the thermal interpretation.

13.1 Quantum buckets and time-resolved events on a screen

Consider the quantum system consisting of a screen and an external classical electromagnetic field. This is a very good approximation to many experiments, in particular to those where the light is coherent. According to the standard interpretation, the analysis (given, for example, in the quantum optics book by Mandel & Wolf [178, Chapter 9]) of the response of the electrons in the screen to the field gives a Poisson process for the electron emission, at a rate proportional to the intensity of the incident field. This is consistent with what is observed when doing the experiment with coherent light. A local measurement of the parameters of the Poisson process therefore provides a measurement of the intensity of the field.

In this analysis, there is nothing probabilistic or discrete about the field; it is just a term in the Hamiltonian of the system. Thus, according to the standard interpretation, the probabilistic response is, in this case, solely due to the measurement apparatus—the screen, the only quantum system figuring in the analysis. At very low intensity, the electron emissions by the screen become visible event by event, and the pattern reflecting the incident field intensity emerges gradually. Effectively, the screen exhibits what is called **shot noise**: It begins to stutter like a motor when fed with gas at an insufficient rate. The stuttering of the screen cannot be due to discrete eigenvalues of an operator representing the intensity—the only operator appearing in the analysis by Mandel and Wolf is an electron momentum operator coupling to a classical field.

The classical external field discussed so far is of course only an approximation to the quantum electromagnetic field, and was only used to show that the discrete response is due to the detector, and only triggered by the interaction with a field. A field

https://doi.org/10.1515/9783110667387-013

mediating the interaction must be present with sufficient intensity to transmit the energy necessary for the detection events; these are for coherent quantum light independent and Poisson distributed even in a full quantum analysis (given by Mandel & Wolf [178, Section 12.10]). In the case of noncoherent quantum light, only the quantitative details change.

The discrete result appears just because each screen electron makes a very inaccurate random binary measurement of the incident field intensity. Each single spot in the gradually appearing interference pattern is measurable to high accuracy, but this is a high-accuracy measurement of the screen only, not of the field (or its particle content). The low accuracies refer to accuracies of the implied field intensity—namely one unit at the responding position and zero units elsewhere, whereas the true intensity is low, but nonzero everywhere where the high intensity interference pattern would show up.

Accepting Mandel and Wolf's detector analysis, nothing depends on the deterministic nature of the thermal interpretation. But the latter explains (see Section 10.4) why neglecting the environment (here the detector details) results in probabilistic features at all, and causes the electrons to exhibit a binary response—remaining bound or escaping to a macroscopic distance, where the effect can be magnified by a photomultiplier.

In the thermal interpretation, one assumes more generally that a similar stuttering effect appears whenever one measures any classical or quantum field at very low intensity, no matter whether a photon field or an electron field or a silver field or a water field is considered. Such a stuttering effect may be illustrated as follows: We consider measuring the rate of classical water flow into a basin by the number of buckets (of a fixed size) per unit time needed to keep the water at a roughly fixed level of height. As long as there is enough flow, the bucket is very busy and the flow is measured fairly accurately. But at very low rates, it is enough to occasionally take out one bucket full of water, and the bucket number is a poor approximation of the flow rate, unless one takes very long intermittent times.

By the same principle, quantum detectors, such as photocells and Geiger counters, act as **quantum buckets**. The sole fact that one has counters already implies that, whatever they measure, the measurements are forced by construction to be integers. This limits the attainable resolution of what is measured as in the example of the 3-digit counter from Section 10.7. If used to measure continuous flow, the uncertainty is always at least 1/2 in the units used for the counting.

13.2 Particle decay

In the thermal interpretation, currents provide the natural description for chemical reactions, collision processes, and particle decay, using the general picture justified in Section 11.7: discrete events emerge from coarse-graining through dissipation together

with the discrete basin structure of the slow manifold of a physical system. In this section, we look at how currents may be used to visualize particle decays.

We explain the principle by considering a particle decay $A \to B + C$, such as $\pi^+ \to \mu^+ + \nu_\mu$. Note that at present, this only gives an intuitive picture of what should happen. The details of this thermal interpretation picture are still conjectural and need to be justified by future analysis of specific models.

At each time t one has three operator-valued effective 4-currents, one for each possibly flowing substance A, B, C. When the center of the reaction is at the origin, the reaction $A \to B + C$ proceeds as follows: At large negative times, the A-density (q-expectation of the time component of the 4-current) is concentrated along the negative z-axis, and the A-current (q-expectation of the 3-vector of space components of the 4-current) is concentrated along the positive z-axis; the B-current and the C-current essentially vanish.

If the reaction happened (which depends on the details of the environment) then, at large positive times, the A-current is negligible, the B-density and C-density are concentrated along two (slightly diverging) rays emanating from the origin in such a way that momentum conservation holds, and the B-current and C-current are concentrated along these rays, too. Otherwise, at large positive times, the A-density is concentrated along the positive z-axis, and the A-current is concentrated along the positive z-axis, too, and the B-current and the C-current remain negligible. During the reaction time, that is, when the fields are concentrated near the origin, one can interpolate the asymptotic happening in an appropriate way. The details are defined by the interaction.

The manifold of slow modes splits into a basin corresponding to the decayed state (with two continuous angle parameters labeling the possible modes), and a basin corresponding to the undecayed state. The metastable transition state at time zero determines, together with the environmental fluctuations, which basin is chosen and which direction is taken. This is comparable to what happens to bending a classical thin iron bar through longitudinal pressure in a random direction, though in that case, the bar must bend, so that there is only one basin, with modes labeled by a single angle. In both cases, one of the continuous labels appears due to the rotational symmetry of the setting around the z-axis. In the case of the decay reaction, the second continuous label arises through another, infinitesimal symmetry at the saddle point at the origin.

This is one of the possible scenarios, probably valid whe the decay happens inside a dense medium (a secondary decay in a bubble chamber, say).

A second scenario applies for a collision experiment in vacuum. Now there is not enough environmental interaction near zero, and after reaching the collision region, the B-current and the C-current should, in case a reaction happens, then take a rotationally symmetric shape. In this case, the path-like particle nature appears only later when the spherical fields reach a detector. The metastability of the detector forces

the two spherical fields to concentrate along two paths, and momentum conservation makes these paths lie weighted-symmetric to the z-axis (geometrically symmetric when the decay products have equal mass). The details are essentially those reported in the 1929 paper by MOTT [187].

In both scenarios, the detection process creates the seeming particle nature of the observation record; see the discussion in Section 12.2.

All **chemical reactions**, **nuclear reactions**, and other collision processes may be treated in a similar way as particle decay. For example, the Compton effect corresponds to a collision process $e + \gamma \rightleftharpoons e + \gamma$ for an electron and a photon. It is a special case of the process $A + B \rightleftharpoons A + B$, to be modeled by the currents of two beams that meet in the collision region.

13.3 The Stern–Gerlach experiment in terms of currents

We now consider the Stern–Gerlach experiment, one of the standard textbook examples used in the context of introducing Born's rule. Here a silver beam is split by a magnetic field into two beams. These beams are observed to produce two spots of silver deposit on a screen.

What can be accurately measured are the positions of the spots in the Stern–Gerlach experiment. That these spots mean an accurate spin measurement is already an interpretation—the traditionally accepted one. It is this interpretation that the thermal interpretation calls into question. It replaces it by the claim that it is an inaccurate measurement of a continuous particle spin with an error of the order of $O(\hbar)$ (as expected for nonclassical measurements, with the correct classical limit). This error is magnified by the experimental arrangement to a macroscopic size.

With the thermal interpretation, we may interpret this experiment either on the level of quantum field theory in terms of currents or by considering individual silver atoms in the beam. The former is the fundamental level and is treated in Section 13.3. The latter is approximate but elementary and is treated in Section 13.4.

In the traditional analysis of the Stern–Gerlach experiment in terms of single silver particles, the dynamics is treated semiclassically for simplicity, and two beams appear as the only possible pathways.

In a field theoretic treatment, the beam is not interpreted classically, but as a quantum field.[1] Thus, the silver is treated as an effective spinor field. It is not a free field, because of the magnetic field in the experiment. The magnetic field is (in the usual semiclassical treatment) a term in the Hamiltonian of the field theory that

1 This is the difference to Schrödinger's failed early attempts to give a continuum interpretation of quantum mechanics in terms of classical fields.

changes the dynamics. It treats different components of the spinor field representing silver in opposite ways, turning a single beam at the source into two while passing the magnet.

Section 7.4 applies, except that the electric current is replaced by a silver current (which means that the formula defining it is a complicated multibody current). To get the total amount of silver deposited, one also needs to integrate over the time of the experiment. Thus, the effective support of the silver current operator $j(x)$ is initially along a single beam, which, upon entering the magnetic field, splits into two beams. The current flows along the direction of the two beams. The amount of silver on the screen at the end measures the integrated beam intensity, the total transported mass. This is in complete analogy to the qubit treated in Section 8.6. Particles need not be invoked.

The intensity of the silver flow is the function of the position on the screen defined by the q-expectation of the incident current integrated over a spot centered at this position. Given the setup, the intensity is positive at the two spots predicted by the mathematics of the theory, and zero elsewhere.

The density operator is that of the whole universe, and the integration in (7.5) is effectively over a cell, to which a piece of the equipment responds, done after the trace computation. The operation $\mathrm{Tr}\rho j(x)$ yields a current J that is nonzero only at two small spots of any cross section (for example, on the screen), and integrating over each spot gives in the symmetric case a total intensity of half of the original beam (before the apparatus) in each spot. Integrating over other regions of the screen gives zero since the integrand is zero there. This is why the silver flows into these two spots, and nowhere else.

Thus, when firing a continuous beam of high intensity one sees two spots, both appearing at essentially the same time. What is measured by a spot is the intensity of the silver flow into the spot, not the spin of single electrons.[2]

In the very low-intensity case, the stuttering effect discussed in Section 13.1 for the double-slit experiment becomes visible at a screen of sufficiently high resolution, and the response of the screen becomes erratic. In particular, if a beam contains only a single particle, the quantum field representing the beam is in a state with sharp particle number $N = 1$, but otherwise nothing changes. Conservation of mass, together with the instability of macroscopic superpositions and randomly broken symmetry forces that only one of the two spots gets marked by a silver atom, just as a classical bar under vertical pressure will bend into only one direction. It is not clear how Nature achieves the former, but this lack of explanation is common to all interpretations of quantum physics.

2 The original Stern–Gerlach paper (and the early discussion about it) indeed talked about "Richtungsquantelung" (quantization of directions) and not of spin measurement. The fact that two beams appear is a consequence of the spin of the electron field, but has nothing per se to do with measuring an electronic spin state. The latter is defined only for single electrons, not for the electron field.

We may interpret the stuttering in terms of the quantum bucket picture from Section 13.1. We may think of each of the two spots on the screen as a quantum bucket measuring the impinging silver flow by counting buckets. We combine the counts into a single pointer variable x by counting left spot events downwards (–1) and right spot events upwards (+1). Each single atom deposited somewhere on the screen is one bucket event reducing the intensity of the inflowing bilocal silver field. It approximates the true value in $[-1, 1]$ (the q-expectation of x) by either +1 or –1, the only possible bucket results. This holds for every single atom, and hence for all the silver that arrives in the two spots. For simplicity, we assume that the silver source is prepared in a state, where the q-expectation of x vanishes. Taking the single buckets as measurement then results each time in a binary measurement of the true (theoretically predicted) uncertain number 0 with uncertainty 1, consistent with a measurement error of 1 in each case. This is completely independent of the flow rate.

13.4 The Stern–Gerlach experiment in terms of particles

An dem Tatbestand, die Elektronenschwärme betreffend, wie er bisher beschrieben wurde, ist nichts Paradoxes. Statt vom Schwarm spreche ich in Zukunft vom einzelnen Elektron und demgemäß von Wahrscheinlichkeit statt von Häufigkeit. Etwas Paradoxes liegt erst in der Aussage, daß σ_x die Komponente eines gewissen Vektors, des Impulsmomentes, in bezug auf die x-Richtung ist. Denn dies involviert doch, wenn wir ein rechtwinkliges Koordinatensystem x y z im Raume einführen und die willkürliche Richtung r die Richtungskosinus a, b, c hat, die Gleichung

$$\sigma_r = a\sigma_x + b\sigma_y + c\sigma_z.$$

Wie verträgt sich das mit dem Umstand, daß σ_r so gut wie $\sigma_x, \sigma_y, \sigma_z$ nur der Werte ±1 fähig ist?

Hermann Weyl, 1927 [300, pp. 8f]

We now consider the Stern–Gerlach experiment not in terms of a field measurement, but as a spin measurement experiment of single silver atoms in the beam. In this case one must—like in every introductory text—treat the silver source as producing an ensemble of single atoms and, ignoring efficiency considerations, assume that each silver atom produces a tiny dot at one of the two spots on the screen.

By a similar analysis as that in Section 10.4, the thermal interpretation would find, looking at the reduced dynamics of the relevant macroscopic q-expectations, that due to the reduced dynamics of the pointer variable (here the relative position of the condensed silver atom), all but the positions at the two spots are unstable, so that the total system is bistable. The essence thus happens at the two spots of the screen. Each silver atom materializes at one of the two spots: It is driven by the interaction with the screen to a transformation of the metastable superposition state before it reaches the screen into a stable state of fairly definite position. The dissipative bistability of the system consisting of a silver atom and the screen effectively acts as discretization process.

In general, due to the representation theory of the compact rotation group, the response of a measurement device to a continuous angular momentum signal, represented by the q-expectation $\langle J\rangle$ of the vector-valued angular momentum J of the measured particle, is an integral multiple of $\frac{1}{2}\hbar$, smeared a little[3] due to the limited precision of the measuring device.

For a high-precision measurement, the angular momentum is essentially discretized to exact multiples of $\frac{1}{2}\hbar$, resulting in discretization errors of order $O(\hbar)$. According to the thermal interpretation, this discretization is an effect of the stochastic dissipativity of the system, consisting of a silver atom and the screen. If we compare the silver atom to a numerical computation and the screen to a floating-point processor, these perturbations are of the same kind as rounding errors in floating-point computations, where the processor forces all unrepresentable real numbers into a representation by one of the two closest machine numbers.

For general angular momentum, $O(\hbar)$ is a tiny amount.[4] But for microscopic measurements at low quantum numbers, these systematic perturbations appear to be large in relative terms.

In particular, the Stern–Gerlach experiment is usually depicted as a model for observing the intrinsic angular momentum of silver atoms. In this case, this perturbation is of the same order as the size of each component of $\langle J\rangle$, which is bounded itself by $\frac{1}{2}\hbar$. Since any number of order $O(\hbar)$ is an $O(\hbar)$ perturbation of any other number of order $O(\hbar)$, no significant difference between the convention of the thermal interpretation and tradition is visible. But since the measurements are at noise level of the stochastic process observed, relative differences of more than 100 percent may appear.

This becomes conspicuous when we change the scale and consider microscopic units. Ignoring the factor of $\frac{1}{2}\hbar$, we represent the spin measurement as a measurement of the q-expectations $\langle\sigma_3\rangle \in [-1,1]$ by means of a binary measurement of the spot on which an arriving silver atom is located, with possible values left spot (–1) or right spot (+1). In the thermal interpretation, each single dot on the screen, at either the left spot (–1) or the right spot (+1), is viewed as an approximate measurement of the q-expectation, which lies somewhere in $[-1,1]$. This approximation is very poor. For

3 In the past, this experiment (and others) could be used for precision measurements of $\hbar$. But from May 20, 2019 onwards, $\hbar$ has—by convention—a fixed (but irrational) value, as part of the 2019 redefinition of SI base units [57]. From then on, one can get (by calibration) exact multiples of $\frac{1}{2}\hbar$, as claimed in Born's rule. However, the thermal interpretation asserts that, since the measurement results are not reproducible, this seeming exactness of the angular momentum measurement is a spurious artifact of measuring it with a quantum bucket.

4 In many cases, classical descriptions can be obtained as limiting cases when Planck's constant $\hbar$ can be set to zero without significant loss of quality of the resulting models. The measurement errors due to the discreteness of spectra (energy level spacings, angular momentum level spacings) are $O(\hbar)$ and vanish in the classical limit. When relevant level spacings become large (photochemistry) excited states behave like different species of particles undergoing something akin to a chemical reaction when a transition happens.

example, when the initial state of the silver atoms is such that its q-expectation is $\langle\sigma_3\rangle = 0$, the error of both binary measurement results ± 1 is 1, but with random signs, consistent with the computed uncertainty, which is also 1.

To improve the accuracy one needs to average over multiple measurements, and gets better results that converge to the true value 0 as the sample size gets arbitrarily large. To see this, one must consider a different operator, namely the mean spin $s = N^{-1}(s_1 + \cdots + s_N)$, where s_k is the σ_3 of the kth silver atom in the ensemble measured. This mean spin operator has an associated (theoretically predicted) uncertain value of $\overline{s} \pm \sigma_s = 0 \pm N^{-1/2}$, which is approximately measured by the mean of the bucket results. This mean is for large N distributed as a Gaussian with zero mean and standard deviation $N^{-1/2}$, matching the prediction.

Born's statistical interpretation treats the measured position of each individual silver atom instead as an exact measurement of the discrete value ± 1 of the corresponding atom, with random signs. Although each single measurement is deemed error-free, the statistical uncertainty resulting from this randomness is still 1.

Clearly, both interpretations account for the same experimental facts, but in different ways. They make very different assumptions concerning the nature of what is to be regarded as the idealized measurement result, to which the actual result is to be compared.

13.5 Relativistic causality

> *Phrases often found in the physical literature as 'disturbance of phenomena by observation' or 'creation of physical attributes of objects by measurements' represent a use of words like 'phenomena' and 'observation' as well as 'attribute' and 'measurement' which is hardly compatible with common usage and practical definition and, therefore, is apt to cause confusion. As a more appropriate way of expression, one may strongly advocate limitation of the use of the word phenomenon to refer exclusively to observations obtained under specified circumstances, including an account of the whole experiment.*
>
> Niels Bohr, 1948 [40]

> *In each experiment, irrespective of its history, there is only one quantum system, which may consist of several particles or other subsystems, created or annihilated at the various interventions.*
>
> Asher Peres and Daniel Terno, 2002 [235, p. 98]

We now consider relativistic causality in a Minkowski spacetime. By working with charts, everything said generalizes to the case of curved spacetime. Quantized spacetime is not discussed, since there is no accepted framework for it.

A **point object** has, at any given time in any observer's frame, properties only at a single point, namely the point in the intersection of its world line and the spacelike hyperplane orthogonal to the observer's 4-momentum at the time (in the observer frame) under discussion. An **extended object** has properties that depend on more than one spacelike-separated spacetime position. A **joint property** is a property that explicitly

depends on more than one spacetime location within the spacetime region swept out by the extended object in the course of time.

Note that, from a fundamental point of view, there is no clear demarcation line that would tell when a system of particles (for example, a molecule, or a solar system) should or should not be regarded as a single object. The thermal interpretation, therefore, treats arbitrary subsystems of a large system as a single object if they behave in some respect like a unity.

We may distinguish three Poincaré invariant definitions of causality.

- **Point causality:** Properties of a point object depend only on its closed past cones, and can influence only its closed future cones. This is used in special relativity, which discusses the motion of a single classical particle in a classical external field.
- **Separable causality:** Joint properties of an extended object consist of the combination of properties of their constituent points. This is intuitively assumed in all discussions of Bell-type nonlocality, and is in conflict with experiments involving highly entangled photons.
- **Extended causality:** Joint properties of an extended object depend only on the union of the closed past cones of their constituent parts, and can influence only the union of the closed future cones of their constituent parts. This is the version that can probably be derived from relativistic quantum field theory, where particles are localized excitations of the quantum field, and hence extended objects.

All three notions of causality agree on the causality properties of point objects ("point causality") but differ on the causality properties of extended objects. If one regards an entangled quantum system as a system of point particles, one runs into lots of counterintuitive conceptual problems. If one regards an entangled quantum system as a single extended system in the above sense, all such difficulties disappear.

Extended causality is the form of causality appropriate for the thermal interpretation. It takes into account what was known almost from the outset of modern quantum physics: quantum objects are intrinsically extended and must be treated as whole.

The extended system view gives the appropriate intuition. The violation of Bell inequalities in experiments, such as those by Aspect [17] (see Section 13.6 below) shows that neither point causality nor separable causality can be realized in Nature. But extended causality is not ruled out by current experiments. Eberhard & Ross [75] gives a proof of causality from relativistic quantum field theory, in the sense that no faster than light communication is possible.

13.6 Nonlocal correlations and conditional information

In this section, we briefly discuss some experimental aspects of nonlocality from the point of view of the thermal interpretation.

Experiments that prove that Bell inequalities are violated imply, together with Bell's theorem, that reality modeled by deterministic process variables is intrinsically nonlocal. The thermal interpretation is based on the deterministic Ehrenfest dynamics of the collection of all q-expectations of the universe. It couples local q-expectations (currents, for example, beam densities and velocities, and field values, for example, idealized pointer readings) to multilocal q-expectations (n-point functions), and accounts in this way for the nonclassical correlations observed in long-distance entanglement, in agreement with Bell's theorem.

A subsystem of a composite system is selected by picking a vector space of quantities (linear operators) relevant to the subsystem. The existence of multilocal q-expectations implies that a composite system is more than its parts. Regarding a tensor product of two systems as two separate subsystems (as often done informally) is appropriate only[5] when all quantities that correlate the two systems are deemed irrelevant.

According to the thermal interpretation, all quantum objects have an uncertain, not sharply definable position. Hence they are intrinsically nonlocal, but usually only slightly. However, two subsystems prepared such that long-distance entanglement is present, must be treated as a whole since their preparation leads to significant multilocality. These composite systems are then very extended even when they are concentrated along two narrow beams.

Consider the example of an entangled 2-photon state. In quantum physics, there is a definite concept of a system in a 2-photon state, but only a fuzzy one of "two photons". Attempting to literally interpret an entangled 2-photon state as consisting of two separate photons and nothing more leads to well-known seemingly paradoxical situations. The correct picture is to consider the system as an extended whole!

Given the quantum mechanical 2-photon system together with the Schrödinger dynamics determined by the associated dispersion relation, Born's rule makes assertions about measurements anywhere in the universe at any future time! Something more nonlocal cannot be conceived. It is, therefore, no surprise that intuition is violated. However, the thermal interpretation renounces the universal validity of Born's rule. Instead, the nonlocal correlations get their natural explanation in terms of extended causality and conditional information.

Consider the conceptual setting of a typical Bell-type experiment, where entangled 2-photon systems are subjected to measurements by two far away observers conventionally called Alice and Bob, at times and positions corresponding to spacelike distances. Their measurement results are later compared by another observer called Charles, say.

In an experiment checking Bell inequalities, an object in the form of a 2-photon system may be prepared at the source by parametric down-conversion, propagating

5 If this is not the case, thinking of the composite system only in terms of its subsystems produces the weird features visible in many discussions of quantum entanglement aimed at the general public.

freely in two opposite directions. Whatever Alice and Bob measure far away depends on the whole 2-photon system. According to the thermal interpretation, the object described by this 2-photon system is an extended object. Over long distances, the uncertainty intrinsic to the 2-photon system becomes huge. The object becomes vastly extended—so nonlocal that the assumptions in Bell's argument are obviously violated. It is not surprising that the conclusions can be violated, too.

Any meaningful use of the notion of causality depends on a notion of before and after, which does not exist in the case of correlations at spacelike distances. Because of the Lorentz invariance of all relativistic arguments, one cannot say that Alice's actions and observations cause (or affect) Bob's observations to be correlated once—as usually assumed—Alice's and Bob's position are causally unrelated. For in this case, there are always Lorentz frames, in which Alice acts later than Bob observes, and others, in which Bob acts later than Alice. So neither can be said to cause (or affect) the other observer (see Peres & Terno [235, 236]). Whatever statistics can be made (by Bob or Charles) from data collected by Bob before Alice's choices or results become available to Charles—it will be completely unaffected by the behavior of Alice and her detector.

But something else from Alice becomes known to Bob faster than light, conditional information. **Conditional information** is information deduced from what—given past and present observations available to a local observer—is known from theory, but is not observed itself by this observer. Additional observations may make conditional information more precise. But as long as part of the data in the condition is not yet known, nothing conclusive is known. Having about tomorrow's weather the conditional information that "Should there be no clouds it will not rain" tells, in fact, nothing useful about the weather tomorrow, unless we have information about tomorrow's clouds.

Similarly, Bob gets conditional information of the kind: "Should Alice have measured X then her result was Y". Because Bob does not know whether the hypothesis holds, he knows nothing useful. Bob's claimed knowledge about the results of Alice's measurement is sound only if Alice actually measured something. If she instead took a nap, or if her detector failed because of a power outage, Bob concluded something wrongly.

Causality only demands that information flow is limited by the speed of light. Nothing in relativity forbids conditional information to be passed faster than light. For example, we know lots of conditional information about what can or cannot happen inside black holes although no information can flow out from there. Such conditional information is obtained from theory independent of observation. *But theoretical conclusions apply instantaneously and have no speed limit.*

In Bell-type experiments, the conditional information and the correlations become actual only when someone (like Charles) has access to the actual data resolving the condition.

But how is it possible that, as in actual long-distance entanglement experiments performed, when Charles checks the findings he finds Alice's conditional informa-

tion unconditionally satisfied? It is easily seen that extended causality is observed. Thus there is no logical consistency problem with special relativity. We can even say a little more: When preparing actual long-distance entanglement experiments, the experimenters have to make sure that nothing will interrupt the expected flow of events needed for a correct experimental performance. This means that they in fact prepare not only the source of the entangled photon pairs, but also Alice's and Bob's environment and the whole environment the photons traverse during the experiment. This preparation must be careful enough to exclude all events that would cause unexpected changes to the intended protocol, or remove the delicate entanglement. But this means that the preparation deposits in the past light cones of both Alice and Bob enough correlated classical information that influences the present of Alice and Bob when they perform their choices. This casts some doubt on the experimental validity of the stringent locality assumptions made in derivations of Bell-type inequalities.

14 A critique of Born's rule

Traditionally, *some* version of Born's rule is always considered to be an indispensable part of any interpretation of quantum mechanics, either as a postulate, or as a result derived—not always on the basis of convincing reasoning—from other postulates. In this section, we have a close look at the possible forms of Born's rule and discuss the limits of its validity.

All traditional foundations of quantum mechanics heavily depend on the concept of (hypothetical, idealized) experiments—far too heavily. This is one of the reasons why these foundations are still unsettled, over 90 years after the discovery of the basic equations for modern quantum mechanics. No other theory has such controversial foundations.

The main reason is that the starting point of the usual interpretations is an idealization of the measurement process that is taken too seriously, namely as the indisputable truth about everything measured. But in reality, this idealization is only a didactical trick for the newcomer to make the formal definitions of quantum mechanics a bit easier to swallow. Except in a few very simple cases, it is too far removed from experimental practice to tell much about real measurements, and hence about how quantum physics is used in real applications.

In experimental physics, measurement is a very complex thing, far more complex than Born's rule (the usual starting point) suggests. To measure the distance between two galaxies, the mass of the top quark, or the Lamb shift, just to mention three basic examples, cannot be captured by the idealistic measurement concept used there, nor by any of the refinements of it discussed in the literature.

In each of the three cases mentioned, one assembles a lot of auxiliary information and ultimately calculates the measurement result from a best fit of a model to the data. Clearly, the theory must already be in place in order to do that. We do not even know what a top quark should be, whose mass we are measuring, unless we have a theory that tells us this!

The two most accurately determined observables in the history of quantum physics, namely the anomalous magnetic moment of the electron and Lamb shift, are not even q-observables!

To present the stage for the criticism of Born's rule in Section 14.3, we first need to clarify the meaning of the term "Born's rule". To distinguish different useful meanings, we look in Sections 14.1–14.2 at the early history of Born's rule.

https://doi.org/10.1515/9783110667387-014

14.1 Early, measurement-free formulations of Born's rule

It is interesting to consider the genesis of Born's rule,[1] based on the early papers of the pioneers of quantum mechanics. This and the next section benefited considerably from discussions with Francois Ziegler, though his view of the history is somewhat different (see ZIEGLER [317]).

The two 1926 papers by BORN [43, 44] (the first being a summary of the second) introduced the probabilistic interpretation that earned Born the 1954 Nobel prize. Born's 1926 formulation[2] *"gives the probability for the electron, arriving from the z-direction, to be thrown out into the direction designated by the angles α, β, γ, with the phase change δ"* does not depend on anything being measured, let alone to be assigned a precise numerical measurement value! Instead it sounds like talk about objective properties of electrons ("being thrown out") independent of measurement. Thus, Born originally did not relate his interpretation to measurement, but to objective properties of scattering processes, no matter whether these were observed.

Rephrased in modern terminology (Born did not have the concept of an S-matrix). Born's statement above is made precise (and generalized) by the following rule:

Born's rule (scattering form): In a scattering experiment described by the S-matrix S,

$$\Pr(\psi_{\text{out}}|\psi_{\text{in}}) := |\psi_{\text{out}}^* S \psi_{\text{in}}|^2$$

is the conditional probability density that scattering of particles prepared in the in-state ψ_{in} results in particles in the out-state ψ_{out}. Here the in- and out-states are asymptotic eigenstates of total momentum, labelled by a maximal collection of independent quantum numbers (including particle momenta and spins).

The scattering form of Born's rule is impeccable and remains until today the basis of the interpretation of S-matrix elements, computed from quantum mechanics or quantum field theory.

The 1927 paper by BORN [45] extends this rule on pp. 173 to probabilities for quantum jumps ("Quantensprung", pp. 172) between energy eigenstates, given by the absolute squares of inner products of the corresponding eigenstates, still using objective rather than measurement-based language[3]: For a system initially in state n given by

---

1 Apparently named such first in 1934 by BAUER [28, p. 302] (*"la règle de Born"*). Before that, during the gestation period of finding the right level of generalization and interpretation, the pioneers talked more vaguely about Born's interpretation of quantum mechanics (or of the wave function). For example, JORDAN [151, pp. 811] writes about *"Born's Deutung der Lösung [der] Schrödingergleichung"*.

2 German original, BORN [43, p. 865f]: *"bestimmt die Wahrscheinlichkeit dafür, daß das aus der z-Richtung kommende Elektron in die durch α, β, γ bestimmte Richtung (und mit einer Phasenänderung δ) geworfen wird"*.

3 *"Das Quadrat $|b_{nm}|^2$ ist gemäß unserer Grundhypothese die Wahrscheinlichkeit dafür, daß das System sich nach Ablauf der Störung im Zustand m befindet."*

Born's equation (9), *"the square $|b_{nm}|^2$ is according to our basic hypothesis the probability for the system to be in state m after completion of the interaction"*. Here state n is the nth stationary state (eigenstate with a time-dependent harmonic phase) of the Hamiltonian.

Born derives this rule from two assumptions. The first assumption, made on pp. 170 and repeated on pp. 171 after equation (5), is that an atomic system is always in a definite stationary state[4]: "*Thus we shall preserve the picture of Bohr that an atomic system is always in a unique stationary state. [...] but in general we shall know in any moment only that, based on the prior history and the physical conditions present, there is a certain probability that the atom is in the nth state.*"

Thus, for the early Born, the objective properties of a quantum mechanical system are the quantum numbers of the (proper or improper) stationary states of the system. This assumption works indeed for equilibrium quantum statistical mechanics, where expectations are defined in terms of the partition function and a probability distribution over the stationary states. It also works for nondegenerate quantum scattering theory, where only asymptotic states figure. However, it has problems in the presence of degeneracy, where only the eigenspaces, but not the stationary states themselves, have well-defined quantum numbers. Indeed, [45] assumes—on pp. 159, remark after his (2) and his Footnote 2—that the Hamiltonian has a nondegenerate, discrete spectrum.

Born's second assumption is his basic hypothesis on pp. 171 for probabilities for being (objectively) in a stationary state[5]: *"there is a certain probability that the atom is in the nth state. We now claim that as measure for this probability of state, one must choose the quantity $|c_n|^2 = |\int \psi(x,t)\psi_n^*(x)dx|^2$"*.

The 1927 paper by Jordan [151, pp. 811] (citing Pauli) extends Born's second assumption further to an objective, measurement independent probability interpretation of inner products (probability amplitudes) of eigenstates of two arbitrary operators, seemingly without being aware of the conceptual problem this objective view poses when applied to noncommuting operators.

The 1927 paper by Pauli [231, pp. 83, Footnote 1] contains the first formal statement of a probability interpretation for position[6]: *"We shall interpret this function in the spirit of Born's view of the "Gespensterfeld" in* [43, 44] *as follows: $|\psi(q_1 \dots q_f)|^2 dq_1 \dots dq_f$*

4 *"Wir werden also an dem Bohrschen Bilde festhalten, daß ein atomares System stets nur in einem stationären Zustand ist. [...] im allgemeinen aber werden wir in einem Augenblick nur wissen, daß auf Grund der Vorgeschichte und der bestehenden physikalischen Bedingungen eine gewisse Wahrscheinlichkeit dafür besteht, daß das Atom im n-ten Zustand ist."*

5 *"[...] eine gewisse Wahrscheinlichkeit dafür besteht, daß das Atom im n-ten Zustand ist. Wir behaupten nun, daß als Maß dieser Zustandswahrscheinlichkeit die Größe $|c_n|^2 = |\int \psi(x,t)\psi_n^*(x)dx|^2$ zu wählen ist."*

6 *"Wir wollen diese [...] Funktion im Sinne der von Born in seiner Stoßmechanik [here he cites* [43, 44]*] vertretenen Auffassung des "Gespensterfeldes" folgendermaßen deuten: Es ist $|\psi(q_1 \dots q_f)|^2 dq_1 \dots dq_f$*

is the probability that, in the named quantum state of the system, these coordinates lie simultaneously in the named volume element $dq_1 \dots dq_f$ of position space." Apart from its objective formulation (no reference to measurement), this is a special case of the universal formulation of Born's rule given below:

The 1927 paper by VON NEUMANN [215, pp. 45] generalizes this statement to arbitrary self-adjoint operators, again stated as an objective (that is, measurement independent) interpretation. For discrete energy spectra and their energy levels, we still read pp. 48: *"unquantized states are impossible" ("nicht gequantelte Zustände sind unmöglich")*.

Note that like Born, Jordan and von Neumann both talk about objective properties of the system independent of measurement. But unlike Born, who ties these properties to the stationary state representation in which momentum and energy act diagonally, Pauli ties it to the position representation, where position acts diagonally, and von Neumann allows it for arbitrary systems of commuting self-adjoint operators.

From either Born's or Jordan's statement, one can easily obtain the following, basis-independent form of Born's rule, either for functions A of stationary state labels, or for functions A of position:

Born's rule (objective expectation form): The value of a q-observable corresponding to a self-adjoint Hermitian operator A of a system in the pure state ψ (or the mixed state ρ) equals on average the q-expectation value $\langle A\rangle := \psi^* A\psi$ (that is, $\langle A\rangle := \mathrm{Tr}\,\rho A$).

The first published statement of this kind seems to be in the 1927 paper by LANDAU [170, (4a), (5)]. The interpretational part is in Footnote 2 there, which states that (the formula corresponding in modern notation to $\langle A\rangle := \mathrm{Tr}\,\rho A$) denotes the probability mean ("Wahrscheinlichkeitsmittelwert"). Again there is no reference to measurement.

14.2 Formulations of Born's rule in terms of measurement

As pointed out by WEYL [300, pp. 2], the derivation from Born's and Jordan's statement does not extend to general operators A, due to noncommutativity and the resulting complementarity. Another consequence of this noncommutativity is that Born's stationary state probability interpretation and Pauli's position probability density interpretation cannot both claim objective status. Therefore, later interpretations relate the notion of value of an observable (being in the n-state, or having position r) more directly to measurement.

The 1927 paper by VON NEUMANN [216] notes (on pp. 248) the problems resulting from noncommuting quantities that cannot be observed simultaneously. For a

die Wahrscheinlichkeit dafür, daß im betreffenden Quantenzustand des Systems diese Koordinaten sich zugleich im betreffenden Volumenelement $dq_1 \dots dq_f$ des Lageraums befinden."

theoretical expectation value with natural properties, the necessity of the formula $\langle A\rangle := \mathrm{Tr}\,\rho A$ with Hermitian ρ (his U) of trace 1 is derived axiomatically (on pp. 255). This is abstract mathematical reasoning independent of any relation to measurement, and hence belongs to the formal (uninterpreted) core of quantum physics. However, the motivation for his axioms, and hence their interpretation, is taken from a consideration (on pp. 247) of the measurement of values in an ensemble of systems, taking the expectation to be the ensemble mean of the measured values. Specialized to a uniform ("einheitlich") ensemble of systems in the same completely known (pure) state ψ of norm 1, he then finds (on pp. 258) that $\rho = \psi\psi^*$, giving $\langle A\rangle := \psi^* A\psi$.

In the present terminology, we may phrase von Neumann's interpretation of q-expectation values as follows:

Born's rule (measured expectation form): If a q-observable corresponding to a self-adjoint Hermitian operator A is measured on a system in the pure state ψ (or the mixed state ρ), the results equal, on average, the q-expectation value $\langle A\rangle := \psi^* A\psi$ (that is, $\langle A\rangle := \mathrm{Tr}\,\rho A$).

Note that the q-expectation value has a formal meaning independent of the interpretation; the measured expectation form of Born's rule asserts that measurements result in a random variable, whose expectation agrees with the formal q-expectation. To justify the "equal", the average in question cannot be a sample average (where only an approximate equal results, with an accuracy depending on size and independence of the sample), but must be considered as the theoretical expectation value of the random variable.

On a purist note, we can only take finitely many measurements on a system. But the expectation value of a random variable is insensitive to the result of a finite number of realizations. Thus, in the most stringent sense, the expectation form of Born's rule says nothing at all about measurement. However, the content of the expectation form is roughly the content of the more carefully formulated statement (MI) discussed in Section 8.1, specialized to a pure state. (MI) does not have the defect just mentioned.

More conventionally, Born's rule is phrased in terms of measurement results and their probabilities rather than expectations. As part of Born's rule, it is usually stated (see, for example, [306]) that each result of a measurement of a q-observable exactly equals one of the eigenvalues.

A precise basic form of Born's rule (often augmented by a more controversial collapse statement about the state after a measurement[7] not discussed here) is the following, taken almost verbatim from Wikipedia [306]:

7 For example, in his famous 1930 book, Dirac [70, pp. 49] states: *"The state of the system after the observation must be an eigenstate of [the observable] α, since the result of a measurement of α for this state must be a certainty"*. In the third edition [71, pp. 36], he writes: *"Thus after the first measurement has been made, the system is in an eigenstate of the dynamical variable ξ, the eigenvalue it belongs to being equal to the result of the first measurement. This conclusion must still hold if the second measurement is not actually made. In this way we see that a measurement always causes the system to jump into*

Born's rule (discrete form): If a q-observable corresponding to a self-adjoint Hermitian operator A with discrete spectrum is measured in a system described by a pure state with normalized wave function ψ then

(i) the measured result will be one of the eigenvalues λ of A, and

(ii) the probability of measuring a given eigenvalue λ_i equals $\psi^* P_i \psi$, where P_i is the projection onto the eigenspace of A corresponding to λ_i.

A related statement is claimed to hold for arbitrary spectra with a continuous part, generalizing both the discrete form and the original form.

Born's rule (universal form)[8]: If a q-observable, corresponding to a self-adjoint Hermitian operator A, is measured in a system described by a pure state with normalized wave function ψ then

(i) the measured result will be one of the eigenvalues λ of A, and

(ii) for any open interval Λ of real numbers, the probability of measuring $\lambda \in \Lambda$ equals $\psi^* P(\Lambda)\psi$, where $P(\Lambda)$ is the projection onto the invariant subspace of A, corresponding by the spectral theorem to the spectrum in Λ.

If the measurement result λ_i is an isolated eigenvalue of A, the universal form reduces to the discrete form, since one can take Λ to be an open interval intersecting the spectrum in λ_i only, and in this case, $P(\Lambda) = P_i$.

Using the spectral theorem, it is not difficult to show that the universal form of Born's rule implies the measured expectation form. Conversely, the measured expectation form of Born's rule almost implies the universal form. It fully implies the second part (ii), from which it follows that the first part (i) holds with probability 1 (but not with certainty, as the universal form claims).[9]

---

an eigenstate of the dynamical variable that is being measured, the eigenvalue this eigenstate belongs to being equal to the result of the measurement."

A 2007 source is Schlosshauer [265], who takes the collapse ("jump into an eigenstate") to be part of what he calls the "standard interpretation" of quantum mechanics, but does not count it as part of Born's rule (pp. 35). On the other hand, Landau & Lifschitz [171, Section 7] explicitly remark that the state after the measurement is in general not an eigenstate.

8 The quantum field theory book by Weinberg [294] pays (on pp. 50 (2.1.7)) lip service to the universal form of Born's rule. The only place where Born's rule is used is on pp. 135 (3.4.7), where instead, the scattering form is employed to get the transition rates for scattering processes. Thus, quantum field theory only relies on the scattering form of Born's rule.

9 This is not just hair splitting. The difference between probability 1 and certainty can be seen by noting that the probability that a random number drawn uniformly from $[0, 1]$ is irrational with probability 1, while measurements usually produce rational numbers.

In general, the expectation form certainly allows finitely many exceptions to (i), and hence says, strictly speaking, nothing at all about the possible measurement results. (This is well illustrated in the story "The Metaphysian's Nightmare" by the logician Bertrand Russell [254]: "There is a special department of Hell for students of probability. In this department there are many typewriters and many

Unfortunately, the application of Born's rule to measurement problems in general is highly questionable. Because of the equivalence just mentioned, it is enough to discuss the universal form of Born's rule.

14.3 Limitations of Born's rule

Every age has scoffed at its predecessor, accusing it of having generalised too boldly and too naively.
Henri Poincaré, 1902 [241, p. 140]

We shall assume the energy of any dynamical system to be always an observable.
Paul Dirac, 1930 [70, p. 38]

Das Resultat ist aber merkwürdiger, als es im ersten Augenblick den Anschein hat. Bekanntlich nimmt $\psi^* \psi$ exponentiell mit wachsendem Abstand vom Atomkern ab. Also besteht immer noch eine endliche Wahrscheinlichkeit dafür, das Elektron in sehr weitem Abstand vom Atomkern zu finden.
Werner Heisenberg, 1930 [120, p. 25]

It is not so that the Schrödinger equation (or its matrix mechanics equivalent) must be supplemented by at least a passable interpretation of the wavefunction (or the matrix elements) before the theory is of any use. Heisenberg, Pauli, Schrödinger, Dirac and their colleagues had firmly established the power of the theory by doing a host of intricate and highly successful calculations before the Born interpretation and the Uncertainty Principle had been put forward.
Kurt Gottfried, 1991 [103, p. 35]

Now, the crux of the problem which worries Wigner so much is that the reduction rule appears to be in contradistinction with the time evolution described by Schrödinger's equation. The answer, which was of course well known to Bohr, but has been made formally clear by the Italians [Daneri, Loinger and Prosperi], is that the reduction rule is not an independent axiom, but essentially a thermodynamic effect, and accordingly, only valid to the thermodynamic approximation.
Leon Rosenfeld, 1972 [251]

Though usually stated as universally valid, Born's rule has severe limitations. In the universal form, it neither applies to photodetection nor to the measurement of the total energy, just to mention the most conspicuous misfits. Moreover, equating the results of measurements with exact eigenvalues is very questionable when the latter are irrational or (as in the case of angular momentum) multiples of a not-exactly- known constant of Nature. In addition, real measurements rarely produce exact numbers (as Born's rule would require it) but (see [220]) numbers that are themselves subject to uncertainty. Because of these limitations and the inherent ambiguities in specifying what constitutes a measurement[10] and what qualifies as a measurement result, subsequent derivations can never claim universal validity either.

monkeys. Every time that a monkey walks on a typewriter, it types by chance one of Shakespeare's sonnets.") However, the application to spin measurements and to quantum information theory requires that (i) holds at least for binary spectra. Hence, the expectation form is slightly deficient.

10 This is a highly nontrivial problem in quantum statistical mechanics; see ALLAHVERDYAN et al. [6].

Problems with Born's rule include:

1. Many things physicists measure have no simple interpretation in terms of a Born measurement. Examples include spectral lines and widths, life times of unstable particles, chemical reaction rates, or scattering cross-sections. Often lots of approximate computations are involved.
2. Born's rule does not cover the multitude of situations, where typically only single measurements of a q-observable are made. In particular, Born's rule does not apply to typical macroscopic measurements, whose essentially deterministic predictions are derived from statistical mechanics.
3. Many measurements in quantum optics are POVM measurements [307], that is, described by a positive operator-valued measure; cf. Section 11.5. These follow a different law, of which Born's rule is just a very special case, where the POVM is actually projection-valued. The general POVM law can be derived from Born's rule applied to a fictitious von Neumann experiment in an extended Hilbert space. This shows its consistency with Born's rule, but still disproves the latter for the actual physical states in the physical Hilbert space.
4. At energies below the dissociation threshold (that is, where the spectrum of the Hamiltonian H, the associated q-observable, is discrete), energy measurements of a system almost never yield an exact eigenvalue of H. For example, nobody knows the exact value of the Lamb shift, a difference of eigenvalues of the Hamiltonian of the hydrogen atom; the (reasonably) precise measurement was even worth a Nobel prize (1955 for Willis Lamb). Indeed, the energy levels of most realistic quantum systems are only inaccurately known.
5. In particular, Born's rule does not apply to the total energy of a composite system, according to Dirac one of the key q-observables in quantum physics, since the spectrum is usually very narrowly spaced and precise energy levels are known only for the simplest systems in the simplest approximations. Therefore, Born's rule cannot be used to justify the canonical ensemble formalism of statistical mechanics; it can at best motivate it.
6. The same holds for the measurement of masses of relativistic particles with 4-momentum p, where a measurement never yields an exact eigenvalue of the mass operator $M := \sqrt{p^2}$. Indeed, the masses of most particles are only inaccurately known.
7. When a particle has been prepared in an ion trap (and hence is there with certainty), Born's rule implies a tiny but positive probability that, at an arbitrarily short time afterwards, it is detected a light year away.[11] Already Heisenberg found

11 Indeed, for a single massive particle, Born's rule states that $|\psi(x,t)|^2$ is the probability density for locating at a given time t the particle at a particular position x anywhere in the universe, and the Fourier transform $|\overline{\psi}(p,t)|^2$ is the probability density for locating at a given time t the particle with a particular momentum p. In the present case, the position density has bounded support, so by a basic

this conclusion strange. Clearly, $|\psi(x)|^2$ cannot be the exact probability density for being detected at x.

8. This argument against the exact probability density interpretation of $|\psi(x)|^2$ works even for relativistic particles. A very general rigorous argument for this instantaneous spreading was given by Hegerfeldt [116]. For example, it applies in the multiparticle framework of Keister & Polyzou [154].
9. Hegerfeldt's result also applies rigorously to electrons in quantum electrodynamics (QED). The free QED electron can be described as follows in terms of a nonlocal single-particle Hamiltonian. The single electron sector of renormalized QED including infrared dressing is invariant under Poincaré transformations, since there is no scattering. Its Hilbert space carries a reducible unitary representation of the Poincaré group. The generator of time defines the Hamiltonian H. The resolvent $(E-H)^{-1}$ equals the renormalized electron propagator, and is given by the Kallen-Lehmann formula associated with some continuous mass density (due to infrared dressing effects) whose support extends from the nominal electron mass to infinity.

   The mass spectrum is nondegenerate and has a branch point at the nominal electron mass, where the continuous mass spectrum has a sharp peak. This means that the free QED electron has an additional mass degree of freedom, which formally behaves like an additional momentum degree of freedom. This mass degree of freedom generates the continuous mass spectrum. Therefore the free QED electron is not an elementary particle[12] in the sense of Wigner but a stable **infraparticle**; see Schroer [268] and Buchholz [56, p. 65f]. The details of the mass density are not completely known but the basic structure, including the branch point, of the electron propagator, is discussed in Section II of Appelquist & Carazzone [12].
10. A no-go theorem for exact measurement by Wigner [305] rules out projective measurements of a particle being in a given region, since the corresponding projector does not commute with all additive conserved quantities. See also Ozawa [227] and Araki & Yanase [13].
11. The measurement of quantum fields is not covered by Born's rule. These are q-observables depending on a space-time argument, and one can prepare or measure events at any particular spacetime position at most once. Thus, it is impossible to repeat measurements, and the standard statistical interpretation in terms of sufficiently many identically prepared systems is impossible. In quan-

theorem of harmonic analysis, the momentum density must have unbounded support. This implies the claim.

12 In the quantum mechanical treatment of multi-electronic systems, the mass degree of freedom is generally suppressed, a sensible approximation given the peaked mass spectrum. Indeed, infrared problems are not much addressed in the literature.

tum field theory, the notion of an ensemble can therefore not be understood as an actual repetition by repeated preparation.

In particular, the measurement of quantum fields is not covered by Born's rule in its standard measurement-based forms. This is the reason why (unlike scattering applications) macroscopic applications of quantum field theory never invoke Born's rule. Indeed, Born's rule cannot be strictly true in relativistic quantum field theory at finite times (that is, outside its use in interpreting asymptotic S-matrix elements), and hence not in Nature.

To uphold Born's rule in the nonrelativistic case, one would, in view of points 4–6, have to treat the concept of measurement as that of a fictitious, infinitely precise measurement. In particular, this would exclude a subjective interpretation of measurement results and associated probabilities in terms of the experimenter's knowledge.

Points 7–9 imply that the wave function, and hence the density operator, encode in their basic operational interpretation unrealistic, highly nonlocal information. Thus, nonlocality is explicitly built into the very foundations of quantum mechanics *as conventionally presented*.[13]

We conclude that Born's rule has, like any other statement in physics, its domain of validity but leads to problems when applied outside this domain.[14] From an analysis of many different q-observables and measurement protocols, it seems that the discrete form of Born's rule needs four conditions for its validity. It is valid precisely for measuring q-observables that simultaneously

- have only a discrete spectrum,
- are measured over and over again in identical states (to make sense of the probabilities),
- there are well-separated adjacent eigenvalues, whose differences are significantly larger than the measurement resolution,
- the measured value is adjusted to exactly match the spectrum, which must be known exactly prior to the measurement.

The universal version of Born's rule has similar limitations also when restricted to purely continuous spectra; in this case it seems to be valid only in Born's original scattering form.

13 Processing nonlocal information, it is no surprise that standard quantum mechanics defined by the Schrödinger equation violates the conclusions of Bell-type theorems (see the discussion and references in Neumaier [194]). It already violates their assumptions!

14 Progressing from the Born rule to so-called positive operator valued measures (POVMs) is already a big improvement, commonly used in quantum optics and quantum information theory. These are adequate for measurements in the form of clicks, flashes or events (particle tracks) in scattering experiments, and perhaps only then.
But these still do not cover measurements of energy, or of the Lamb shift, or of particle form factors.

15 Pure states and mixed states

In this chapter, the concepts of states and ensembles are critically reviewed. Among other things, it is shown that if the state of every composite quantum system contains all information that can be known about a system, states cannot be pure states.

In Section 15.1, we criticize the standard practice of treating the state vector (that is, a pure state) as the complete description of a (single, or frequently prepared) quantum system. Section 15.2 shows that the density operator contains objective information about single quantum systems. This conclusion is reinforced in Section 15.3, which discusses Gibbs' in his time revolutionary notion of fictitious ensembles for representing single macroscopic systems in equilibrium. The final Section 15.4 shows that pure states for objects in a bounded region of space are also questionable from the perspective of relativistic quantum field theory.

15.1 What is a state?

In physics, the state of a physical system (whether classical or quantum) gives a complete description of the system at a given time. The following is a concise formulation of this:

(S1) The state of a system (at a given time) encodes everything that can be said (or "can be known") about the system at this time, including the possible predictions for later times, and nothing else.

Thus, every property of the system can (in principle) be computed from its state.

For a complex system, knowledge about the whole system is usually obtained by collecting knowledge about its various parts. This makes sense only if we also require:

(S2) Every property of a subsystem is also a property of the whole system.

Indeed, not knowing something about the subsystem means not knowing everything about the system as a whole, and hence not knowing the precise state of the system.

Therefore, common sense dictates that a sound, observer-independent interpretation of quantum physics should satisfy (S1) and (S2).

Now (S2) says that the state of the full system determines all properties of any of its subsystems. Hence, it determines—by (S1)—the state of each subsystem to the last detail. Thus, we conclude:

(S3) The state of a system determines the state of all its subsystems.

A macroscopic body should therefore have a valid microscopic quantum description—a quantum state—that determines all observable properties on every level. In particular, an approximate hydromechanical classical description for the most important observable properties, namely the q-expectation values of the fields, must be ob-

https://doi.org/10.1515/9783110667387-015

tainable from this exact quantum state. (The process to achieve this is usually called coarse graining.)

Property (S3) must hold for logical reasons even though, in practice, we may never know the precise state of the system and/or the subsystems. Indeed, we usually know only very little information about any system, unless the latter is so tiny that it can be fully described by very few parameters.

Unfortunately, none of the mainstream versions of the interpretation of quantum mechanics (that is, those not invoking hidden variables) are anywhere presented in a form that would satisfy our conclusion (S3). Since the deficiency always has the same root—the treatment of the density operator as representing a state of incomplete knowledge, a statistical mixture of pure states—it is enough to discuss one specific interpretation. We shall look at the interpretation given in the very influential treatise of theoretical physics by Landau & Lifshitz [171, 172]. They start their discussion of quantum mechanics with a particular version of Born's rule:

> [171, p. 6] *"The basis of the mathematical formalism of quantum mechanics lies in the proposition that the state of a system can be described by a definite (in general complex) function $\Psi(q)$ of the coordinates. The square of the modulus of this function determines the probability distribution of the values of the coordinates: $|\Psi^2|dq$ is the probability that a measurement performed on the system will find the values of the coordinates to be in the element dq of configuration space. The function Ψ is called the wave function of the system. [...] If the wave function is known at some initial instant, then, from the very meaning of the concept of complete description of a state, it is in principle determined at every succeeding instant".*

Thus, in terms of our formal core, the complete description of the system is declared by Landau & Lifshitz to be a pure state, and the properties of the system are declared to be the probabilities of potential measurement results. They then consider parts (subsystems) and observe:

> [171, p. 7] *"Let us consider a system composed of two parts, and suppose that the state of this system is given in such a way that each of its parts is completely described.†"*

Footnote: *"† This, of course, means that the state of the whole system is completely described also. However, we emphasize that the converse statement is by no means true: a complete description of the state of the whole system does not in general completely determine the states of its individual parts"*

Therefore, they explicitly deny (S3). In fact, except in the special case discussed in the context of the above quote—where the state factors into a tensor product of states of the subsystems—*they do not indicate at all how the state of a system and that of its parts are related*. More mysteriously, nowhere in the literature seems to be a discussion that would tell us anything on the formal level about the relationship between the pure state of a system and the pure state of a subsystem. There seems to be no such relation, except in the idealized, separable case mentioned above, usually assumed to be valid only before an interaction happens.

But this would mean that the quantum state of a physics lab has nothing to do with the quantum states of the equipment in it, and of the particles probed there! This is *very* strange for a science such as physics that studies large systems primarily by decomposing them into its simple constituents and infers properties of the former from collective properties of the latter.

This truly unacceptable situation shows that there is something deeply wrong with the traditional interpretations.[1] It is no surprise that this leads to counterintuitive paradoxes in situations, such as experiments with entangled photons, where a larger (for example, 2-photon) system is prepared, but its constituents (here 2 single photons) are observed.

15.2 The nature of mixed state?

It is unquestionable that we can never know the exact quantum state of a physics lab or a piece of its equipment. Because of this, it has become respectable to interpret quantum mechanics not in terms of what is, but in terms of what is known to the person modeling a physical system. The system state then becomes a complete description no longer of the physical system, but of the knowledge available.[2] Uncertainty about the pure state is then modeled as a probability distribution for being in a pure state. Averaging with corresponding weight leads to more general mixed states described by density operators. In this context, Landau & Lifshitz write:

> [172, p. 16] *"The quantum-mechanical description based on an incomplete set of data concerning the system is effected by means of what is called a density matrix [...] The incompleteness of the description lies in the fact that the results of various kinds of measurement which can be predicted with a certain probability from a knowledge of the density matrix might be predictable with greater or even complete certainty from a complete set of data for the system, from which its wave function could be derived."*

Based on this, they derive the interpretation of the q-expectation $\langle A\rangle := \operatorname{Tr}\rho A$ as the expectation value of A in a mixed[3] state ρ:

> [172, p. 17] *"The change from the complete to the incomplete quantum-mechanical description of the subsystem may be regarded as a kind of averaging over its various ψ states. [...] the mean value $\overline{f}$ becomes the trace (sum of diagonal elements) of this operator"*

1 In addition, there are the traditional difficulties of interpretations of quantum mechanics, well summarized in Section 3.7 of the quantum mechanics textbook by Weinberg [297].

2 It is usually not addressed whose knowledge this is; presumably it is the knowledge of the person creating the mathematical model of the quantum system. Taken at face value, this would make the system state a function of the mental state of the modeler's mind, another truly unacceptable situation.

3 In the literature, such a mixed state is often called a **proper mixed state**, in contrast to the **improper mixed states** arising as a reduced density operator through a partial trace; cf. Section 10.2.

However, on the next page they call their description an illustration only, denying it any trace of reality:

> [172, p. 18] *"It must be emphasised that the averaging over various ψ states, which we have used in order to illustrate the transition from a complete to an incomplete quantum-mechanical description, has only a very formal significance. In particular, it would be quite incorrect to suppose that the description by means of the density matrix signifies that the subsystem can be in various ψ states with various probabilities and that the averaging is over these probabilities. Such a treatment would be in conflict with the basic principles of quantum mechanics.*
> *The states of a quantum-mechanical system that are described by wave functions are sometimes called pure states, as distinct from mixed states, which are described by a density matrix. Care should, however, be taken not to misunderstand the latter term in the way indicated above.*
> *The averaging by means of the statistical matrix according to (5.4) has a twofold nature. It comprises both the averaging due to the probabilistic nature of the quantum description (even when as complete as possible) and the statistical averaging necessitated by the incompleteness of our information concerning the object considered. For a pure state only the first averaging remains, but in statistical cases both types of averaging are always present. It must be borne in mind, however, that these constituents cannot be separated; the whole averaging procedure is carried out as a single operation, and cannot be represented as the result of successive averagings, one purely quantum-mechanical and the other purely statistical".*

Thus, Landau and Lifschitz reject the subjective, knowledge-based view, as one cannot divide the information contained in the density operator into an objective, pure part corresponding to the objective properties of the system, and a statistical part accounting for the lack of knowledge.

But this also means that their derivation of the interpretation of the q-expectation $\langle A\rangle := \mathrm{Tr}\,\rho A$ as expectation value is invalid, being based on an invalid illustration only. Note that this formula is heavily used in quantum statistical mechanics and quantum field theory. It is often applied there in contexts, where no measurement at all is involved, and when it is not even clear how one should measure the operators in question.[4] Indeed, most of quantum statistical mechanics is not concerned with measurement at all. In all these cases, the connection to measurement and hence to Born's rule is absent, and even the hand-waving "illustrative" derivation given is spurious, as the items going into the derivation are never actually measured.

On the other hand, the use of the density operator is central to quantum statistical mechanics. The fact that the latter predicts qualitatively and quantitatively the thermodynamics of macroscopic systems shows that the density operator contains objective, knowledge-independent information, and is the true carrier of the state information in quantum physics.

This is one of the reasons why the present book features the density operator as basic, both in the description of the formal core of quantum physics presented in

4 We mentioned already in Section 14.3 the problems with interpreting measurements of the energy, corresponding to the operator H figuring in all of quantum statistical mechanics.

Section 2.1 and in the thermal interpretation. Pure states then appear as idealizing approximation under the conditions discussed in Section 2.5.

15.3 What is an ensemble?

> *We may imagine a great number of systems of the same nature, but differing in the configurations and velocities which they have at a given instant, and differing not merely infinitesimally, but it may be so as to embrace every conceivable combination of configuration and velocities. [...] The first inquiries in this field were indeed somewhat narrower in their scope than that which has been mentioned, being applied to the particles of a system, rather than to independent systems.*
>
> Josiah Willard Gibbs, 1902 [96, pp. vii–viii]

> *So aufgefaßt, scheint die Gibbssche Definition geradezu widersinnig. Wie soll eine dem Körper wirklich eignende Größe abhängen nicht von dem Zustand, den er hat, sondern den er möglicherweise haben könnte? [...] Es wird eine Gesamtheit mathematisch fingiert. [...] erscheint es schwierig, wenn nicht ausgeschlossen, dem Begriffe der kanonischen Gesamtheit eine physikalische Bedeutung abzugewinnen.*
>
> Paul Hertz, 1910 [126, pp. 226f]

The thermal interpretation of quantum physics says that, consistent with statistical thermodynamics, a q-expectation (q-ensemble mean) is interpreted as an (in principle) approximately measurable quantity. Except when the statistical context is immediate (such as in computer simulations), the q-expectation should not be interpreted as a statistical average over a population of many realizations.

The q-expectation, conventionally called the ensemble expectation, becomes in the thermal interpretation simply the uncertain value. Therefore, the notion of q-ensemble is to be understood not as an actual repetition by repeated preparation. It should be understood instead in the sense of a fictitious collection of imagined copies, of which *only one* is actually realized, giving an intuitive excuse for using the statistical formalism for a single system.

The association of a fictitious ensemble to *single* thermal systems goes back to Gibbs, the founder of the ensemble approach to classical statistical mechanics. He was very aware that thermodynamics, and hence statistical mechanics, applies to single physical systems. His arguments are today as cogent as when he introduced them.

In classical statistical mechanics, the distinction between the deterministic and stochastic description becomes blurred, as *each single* macroscopic system is already described by a phase space density (multiparticle distribution function), although the latter behaves mathematically in every respect like a probability density that expresses the properties of a population of identical systems.

This tension in the terminology is already visible in the famous statistical mechanics textbook by Gibbs [96] in 1902, where he introduced in the preface (from which the above quote is taken) fictitious ensembles to bridge the conceptual gap.

Thus, to deduce properties of macroscopic materials, Gibbs uses an ensemble of macroscopic systems, in contrast to Boltzmann, who introduced statistical mechanics for gases by using ensembles of microscopic atoms. Treating a collection of particles in a gas as an ensemble (as Boltzmann did) makes statistical sense, as there are a huge number of them. But the Gibbs formalism is applied to single macroscopic systems, such as a brick of iron rather than to its many constituents. Treating a single system as part of a fictitious ensemble was a very bold step taken by Gibbs. This allowed him to extend Boltzmann's work from ideal gases to arbitrary chemical systems, in a very robust way. His statistical mechanics formalism, as encoded in the textbook (Gibbs [96]), survived the quantum revolution almost without change; the book reads almost like a modern book on statistical mechanics!

Though exceedingly successful, Gibbs' fictitious ensemble raised in his time severe objections in the physics community, such as the response by Hertz quoted above, who complained that an ensemble is feigned mathematically. Of course Gibbs was aware that imagined systems have no physical implications, but these were needed at a time when mathematics had not yet the abstract character that it has today.

Today, mathematical theories are simply formal systems used without hesitation in applications, in which the terms may mean something completely different from their meaning in the uses that gave the names to the terms. For example, the mathematical notion of a vector is today an abstract tool routinely used in contexts, where the original geometric notion of a vector is meaningless: No physicist thinks of a quantum mechanical state vector in terms of a little arrow depicting a translational motion.

In the same spirit, mathematical statistics (and hence statistical mechanics) may be used as a tool, in which expectation values figure as abstract notions without the need to imagine an ensemble of copies of the single system under study, over which the expectation would be an imagined average. Thus, we are liberated from having to think of the mathematical q-expectation values manipulated in statistical mechanics as being true averages over fictitious copies without a physical meaning. Instead, the statistical terminology is simply a reminder of which laws (originally stemming from statistical data analysis) are applicable to these values, in the same way as the geometric terminology of a vector indicates the laws valid for manipulating objects behaving algebraically like vectors.

15.4 Pure states in quantum field theory

That pure states cannot have a fundamental meaning can also be seen from the perspective of quantum field theory. It is a very little known fact that, in any interacting relativistic quantum field theory, the notion of a pure state loses its meaning. Results from algebraic quantum field theory (see Yngvason [311, pp. 12]) imply that all local algebras induced by a relativistic quantum field theory on a causal diamond (an intersection of a future cone and a past cone with nonempty interior) are **factors of type**

III_1 in von Neumann's classification of factors as refined by Connes [61]. Picking such a causal diamond containing our present planetary system implies that we may assume the algebra of observables currently accessible to mankind to be such a factor of type III_1. Remarkably, such a C^*-algebra $\mathbb{A}$ has no pure states [311, pp. 14].[5] Therefore, in these representations, one cannot rigorously argue about states by considering partial traces in nonexistent pure states! This shows that pure states must be the result of a major approximating simplification, and not something fundamental.

Note that $\mathbb{A}$ has infinitely many unitarily inequivalent irreducible representations on Hilbert spaces (corresponding to the different superselection sectors of the theory). But in each such representation, the algebra $\mathbb{A}$ of bounded q-observables is vanishingly small compared to the algebra of all bounded operators. A vector state in the Hilbert space $\mathbb{H}$ of an irreducible representation of a local algebra $\mathbb{A}$ of type III_1 (which is a pure state in $\mathbb{H}$) can therefore still be mixed as a state of $\mathbb{A}$.[6] The vector state is guaranteed to be pure only relative to the algebra of all bounded operators on $\mathbb{H}$. But this algebra is far bigger than the algebra $\mathbb{A}$, and contains lots of operators that have no interpretation as q-observables. This is the essential difference to the case of type I algebras, realizable in a Fock space, which have many pure states. These algebras are the local algebras of free quantum field theories, and only encode systems of noninteracting particles.

Therefore, what breaks down in quantum field theory is the simple equation "q-observable = self-adjoint Hermitian linear operator". Once this equation is broken, the question whether a state is pure becomes dependent on the precise specification, of which operators are q-observables. In gauge theories, the situation is further complicated by the fact that the local algebras have a nontrivial center consisting of charges that in each irreducible representation are represented trivially. Thus, a single irreducible representation on a single Hilbert space (corresponding to a single superselection sector) is no longer sufficient to characterize the complete algebra of local q-observables.

To give up the assumption that every bounded self-adjoint operator is a q-observable has serious consequences for the interpretation of quantum physics. Indeed, a test for a pure state is in terms of q-observables an observation of the orthogonal projector to the subspace spanned by the state. If this is not a q-observable,

5 An explicit example of a factor of type III_1, involving an infinite array of spin-1/2 particles, is given in equations (27) and (29) of Yngvason [311].

6 For example, for an arbitrary mixed state of a Hilbert space $\mathbb{H}_0$, the GNS construction produces another Hilbert space, in which this state is pure. Note that in quantum physics, the GNS construction is of limited value only, as this Hilbert space depends on the state one started with, whereas standard quantum mechanics works with pure states contained in a **fixed** Hilbert space. One therefore needs a distinguished state to define a Hilbert space. Now the only distinguished state in quantum field theory is the vacuum state. But for gauge theories, such as quantum electrodynamics, the Hilbert space corresponding to this vacuum representation does not contain any charged state!

then it is (in principle) impossible to make this test. Therefore, testing for being in a pure state is impossible, since these are no longer physical propositions.

So one cannot decide whether a system is in a pure state. To assume this is thus a metaphysical act, and one can dispense with it without any loss of information. But then, all traditional interpretations break down completely, since they start with Born's rule for pure states and derive everything else from it.

16 Traditional interpretations

Über die physikalische Interpretation der Formeln sind die Meinungen geteilt.

Max Born, 1926 [44, p. 803]

Das Einzelsystem trägt wirklich die Fähigkeit in sich, einem bestimmten Meßvorgang gegenüber in verschiedener Weise zu reagieren, d. h. verschiedene Meßwerte für ein und dieselbe Zustandsgröße zu liefern: welchen, hängt vom Zufall ab, oder besser wohl von den Phasenbeziehungen zwischen dem System und dem Meßinstrument.

Erwin Schrödinger, 1929 [64, Vorwort]

I reject the basic idea of contemporary statistical quantum theory, insofar as I do not believe that this fundamental concept will provide a useful basis for the whole of physics.

Albert Einstein, 1949 [80, p. 666]

It is usually believed, that the current orthodox theory actually accounts for the 'nice linear traces' observed in the Wilson chamber etc. I think this is a mistake, it does not.

Erwin Schrödinger, 1958 [270, p. 163]

Personally I still have this prejudice against indeterminacy in basic physics.

Paul Dirac, 1972 [72, p. 7]

When it comes to specifying exact details, one discovers that we cannot rigorously define what quantum mechanical amplitudes are, what it means when it is claimed that 'the universe will collapse with such-and-such probability', what and where the observers are, what they are made of, and so on. Yet such questions are of extreme importance if one wants to check a theory for its self-consistency, by studying unitarity, causality, etc.

Gerard 't Hooft, 1999 [137, p. 95]

My own conclusion (not universally shared) is that today there is no interpretation of quantum mechanics that does not have serious flaws.

Steven Weinberg, 2013 [297, p. 95]

I consider it to be an intellectual scandal that, nearly one hundred years after the discovery of matrix mechanics by Heisenberg, Born, Jordan and Dirac, many or most professional physicists – experimentalists and theorists alike – admit to be confused about the deeper meaning of Quantum Mechanics (QM), or are trying to evade taking a clear standpoint by resorting to agnosticism or to overly abstract formulations of QM that often only add to the confusion. [...] I felt that the subject had better remain a hobby until later in my career. [...] But when I was approaching mandatory retirement I felt an urge to clarify my understanding of some of the subjects I had had to teach to my students for thirty years.

Jürg Fröhlich, 2019 [92, p. 1]

From its very beginning in 1926, what turned out to be the formal core of quantum mechanics had conflicting interpretations—initially the deterministic view of Schrödinger and the statistical view of Born. In 1929, Schrödinger conceded the need

https://doi.org/10.1515/9783110667387-016

for a statistical interpretation. But the details remained controversial. Today, after almost 100 years of interpretation quarrels, the matter is still not settled. As the above quotes show, many of the founders of quantum mechanics were never satisfied with the interpretation of quantum mechanics, and even some of today's Nobel prize winners spent significant effort on the interpretation issue.[1]

In this chapter, we review some of the traditional interpretations from the perspective of the thermal interpretation.

16.1 A classification of interpretations

With regard to the interpretation of quantum physics, this book features primarily the thermal interpretation. But as already mentioned in Section 1.1, a multitude of other interpretations of quantum mechanics exist; most of them in several variants. We distinguish the following types:

(I) **Individual interpretations** (such as certain variants of the Copenhagen interpretation), where the state of a system is determined by an individual realization of the system, and contains the information about everything that can possibly be said about it.

(S) **Statistical interpretations** (such as the minimal interpretation), where the state of a system says (except in special cases) nothing about a single system, but is only about statistical predictions of actual measurements on an ensemble of similarly prepared systems.

(K) **Knowledge interpretations**, where a state says nothing objective about the systems modeled, but is only about the subjective knowledge of these systems.

(O) **Other interpretations**, where a state consists (as in Bohmian mechanics) of more than a state vector or density operator, is (as in many worlds interpretations) by conception about more than actual events recorded in actual experiments, or where (as in objective collapse theories) quantum mechanics is modified by modifying the Schrödinger equation.

In the mainstream interpretations of the types (I) and (S), the result of a single measurement is—in contrast to classical mechanics—not even theoretically determined before the measurement is done.[2]

The thermal interpretation is of type (I), but shares with classical mechanics the feature that the result of every single measurement is fully determined by the state of an isolated system containing the system measured and the detector. As we shall see, the mainstream interpretations may be regarded as partial versions of the thermal interpretation.

1 They did so only after their retirement: While pursuing their career, they had to research issues better rewarded by the scientific community and kept—like Fröhlich and probably most quantum physicists—the foundational issues on the back burner.

2 However, expositions of both views generally prefer to remain vague, or even silent about this.

In particular, certain puzzling features of both the Copenhagen interpretation and the statistical interpretation get their explanation through the thermal interpretation of quantum field theory. We shall see that these peculiar features get their natural justification in the realm for which they were created—the statistics of few particle scattering events.

Interpretations of the types (K) and (O) have little in common with the thermal interpretation and are not further discussed.

16.2 The Copenhagen interpretation

The concept of observation is in so far arbitrary as it depends upon which objects are included in the system to be observed. Ultimately every observation can of course be reduced to our sense perceptions. The circumstance, however, that in interpreting observations use has always to be made of theoretical notions, entails that for every particular case it is a question of convenience at what point the concept of observation involving the quantum postulate with its inherent 'irrationality' is brought in.

Niels Bohr, 1927 [39, p. 580]

Um zur Beobachtung zu gelangen, muss man also irgendwo ein Teilsystem aus der Welt ausschneiden und über dieses Teilsystem eben 'Aussagen' oder 'Beobachtungen' machen. Dadurch zerstört man dort den feinen Zusammenhang der Erscheinungen und an der Stelle, wo wir den Schnitt zwischen dem zu beobachtenden System einerseits, dem Beobachter und seinen Apparaten andererseits machen, müssen wir Schwierigkeiten für unsere Ansschauung erwarten. [...] Jede Beobachtung teilt in gewisser Weise die Welt ein in bekannte und unbekannte oder besser: mehr oder weniger genau bekannte Grössen.

Werner Heisenberg, 1927 [118, pp. 593f]

wir müssen die Welt immer in zwei Teile teilen, der eine ist das beobachtete System, der andere der Beobachter. In der ersteren können wir alle physikalischen Prozesse (prinzipiell wenigstens) beliebig genau verfolgen, in der letzteren ist dies sinnlos. Die Grenze zwischen beiden ist weitgehend willkürlich

John von Neumann, 1932 [217, pp. 223f]

Aus diesem Zwiespalt ergibt sich die Notwendigkeit, bei der Beschreibung atomarer Vorgänge einen Schnitt zu ziehen zwischen den Meßapparaten des Beobachters, die mit den klassischen Begriffen beschrieben werden, und dem Beobachtungsobjekt, dessen Verhalten durch eine Wellenfunktion dargestellt wird. Während nun sowohl auf der einen Seite des Schnittes, die zum Beobachter führt, wie auf der anderen, die den Gegenstand der Beobachtung enthält, alle Zusammenhänge scharf determiniert sind – hier durch die Gesetze der klassischen Physik, dort durch die Differentialgleichungen der Quantenmechanik –, äußert sich die Existenz des Schnittes doch im Auftreten statistischer Zusammenhänge. An der Stelle des Schnittes muß nämlich die Wirkung des Beobachtungsmittels auf den zu beobachtenden Gegenstand als eine teilweise unkontroIlierbare Störung aufgefaßt werden. [...] Entscheidend ist hierbei insbesondere, daß die Lage des Schnittes – d. h. die Frage, welche Gegenstände mit zum Beobachtungsmittel und welche mit zum Beobachtungsobjekt gerechnet werden – für die Formulierung der Naturgesetze gleichgültig ist.

Werner Heisenberg, 1934 [121, pp. 670f]

The **Copenhagen interpretation** is the interpretation of quantum mechanics first expressed in 1927 by Bohr and Heisenberg. Until 1970, it has been (in various variants) the almost generally accepted interpretation, though there is no document defining it precisely; its contents was stated in varying ways depending on the occasion. One of the probable reasons is that it had sufficient definiteness to guide theory, experiment, and their relationship, and was at the same time sufficiently vague that it allowed each user to make sense of its paradoxical features in a personal, subjective way.

In our classification of interpretations of quantum mechanics, the Copenhagen interpretation belongs to type (I); the term 'knowledge' used first by HEISENBERG [117] was not understood in the subjective way used in (K), but as the objective (through thought experiments theoretically accessible) knowledge of what is real and in principle observable about the system, whether observed or not.

One important feature of the Copenhagen interpretation is the so-called **Heisenberg cut**, first described by HEISENBERG [118] and BOHR [39, pp. 580,584]—the artificial splitting of the world into a quantum domain and a classical domain. VON NEUMANN [217, p. 223f] showed that this cut can be placed fairly free without affecting the main conclusions.

Whereas adequate for microscopic systems, the concept of a necessary cut fails systematically for sufficiently large systems. For example, as all measurements are done within the solar system, it excludes treating the solar system as a quantum system (for example, measuring the mass of the earth).

As mentioned already in Section 7.6, the thermal interpretation nowhere imposes a cut between microscopic and macroscopic. It is not needed: A paper by JEON & YAFFE [149] derives the hydrodynamic equations from quantum field theory without assuming a Heisenberg cut. Only the thermal interpretation is (implicitly) invoked, which allows them to identify field expectations with the classical values of fields.

According to the thermal interpretation, classical physics appears gradually as systems become more macroscopic. In continuation of the discussion in Section 11.3, we call a quantum system, whose relevant quantities have a negligible uncertainty, a **classical system**. It is typically described by nonequilibrium thermodynamics, as deduced from quantum statistical mechanics; see Section 7.7. Thus, a classical system is still quantum mechanical when modeled in full detail, but only the macroscopic variables modeled by statistical mechanics are deemed to be relevant. Therefore, the thermal interpretation leads to a gradual change from quantum to classical as the system gets larger, and the uncertainty of the collection of relevant quantities decreases.

But the thermal interpretation realizes a modified version of the Heisenberg cut as the choice of relevant variables in the coarse-grained description, which defines the split between system and environment. According to the thermal interpretation, there is no sharp cut, but a smooth fuzzy boundary, of the same kind as the boundary between the Earth's atmosphere and interplanetary space. The bigger one makes the instrument, the more classical it becomes and the more accurate become the pointer

positions. In place of deciding where to place the Heisenberg cut, we now have to decide at which level of description the $O(N^{-1/2})$ corrections can be neglected. This is a decision just like the decision of whether or not to include into the classical description of a pendulum the surrounding air and the way it is suspended, or whether taking it into account with a damping term, or even neglecting that as well, is enough.

In quantum-classical approximations of a quantum system, the Heisenberg cut is explicitly modeled by allowing for both classical and quantum degrees of freedom, neglecting only irrelevant variables. As we have seen, the quantum-classical description naturally fits the thermal interpretation, since q-expectations occur explicitly in the dynamics.

Die 'Bahn' entsteht erst dadurch, daß wir sie beobachten

Werner Heisenberg, 1927 [117, p. 185]

Another puzzling feature of the Copenhagen interpretation is that an individual few-particle system has **no definite properties** before it is observed. Taking the Copenhagen interpretation as an irreducible description of the nature of things leaves one puzzled regarding how the observing instrument can be informed about what it observes, being virtually nonexistent before the act of observation. In his 1927 paper famous for the uncertainty relation, Heisenberg asserted that the path (of a particle) is created only through the act of observation. Thus, the observation creates the properties. But it must be created by something to be observed! The thermal interpretation gives the natural answer that this happens because the fields provide the information about what is there to cause the detectors to respond, so that something is observed. Hence, when measured, particles appear as detection events created by the detector and mediated by fields (see Section 12.2).

In photodetection, tradition takes the individual detection results too seriously and dogmatically[3] interprets the random counting events as signals of single photons arriving, with all the spooky problems associated with this view. In contrast, the thermal interpretation treats it (as Stokes would have done it in 1852) as a very uncertain measurement of energy density. Then the particle aspect completely disappears. This is an advantage since, as we have seen in Section 12.7, it is difficult to specify, even informally, a particle picture at finite times in terms of the underlying relativistic quantum field description. How to do this with some degree of mathematical precision is an unsolved problem.

Fragen wir also nicht, wo ist ein Teilchen genau, sondern begnügen wir uns, zu wissen, daß es in einem bestimmten größeren Raumteil ist: dann verschwindet der Widerspruch zwischen Wellen und Corpusculartheorie.

Max Born, 1929 [46, p. 116]

3 There is no way to test this assumption empirically.

There is no doubt that while an atom is in an ion trap, it has a definite but uncertain position. We know it is there and can check efficiently the duration of its presence. Indeed, to do experiments with single atoms at all we need to know that they are there! In the Copenhagen interpretation, this knowledge was outside the quantum domain, on the classical side of the Heisenberg cut. The thermal interpretation preserves the reality of atoms being somewhere reasonably well localized while rejecting the idealization (made in classical point mechanics) that this position is given by an exact 3-dimensional real vector. This assumption leads to paradoxes both in classical electrodynamics and in certain quantum mechanics experiments. In the thermal interpretation, it is avoided from the start, since all quantities come with their intrinsic uncertainty.

A "particle trace" on a photograph is also measurable. Tradition postulates that a corresponding particle existed that left this trace. But this statement is not experimentally refutable by any means, hence is a metaphysical assumption. Assuming it, we can **infer** the uncertain position and momentum of a particle that was supposed to create it at the time of its creation. But we might also argue, as in Section 12.2, and declare—in analogy with the bullet experiment discussed there—the tracks as a measure of impact quality, not associated with any particle!

Grassi [107] and Jeon & Heinz [148, Section 5.3] (and many others) treat (in line with the thermal picture) interacting elementary "particles" not as particles, but as quantum fluids; only their freeze-out in scattering experiments produces particle-like objects, in a way more or less analogous to how condensed droplets appear in saturated liquids. However, the quantization introduces a discrete element into the quantum numbers (and hence the number and distribution) of the resulting droplets, measurable as impact events or particle tracks.

> *Jede Ortsbestimmung reduziert also das Wellenpaket wieder auf seine ursprüngliche Größe*
>
> Werner Heisenberg, 1927 [117, p. 186]

> *The state of the system after the observation must be an eigenstate of [the observable] α, since the result of a measurement of α for this state must be a certainty.*
>
> Paul Dirac, 1930 [70, p. 49]

A third significant feature of the Copenhagen interpretation is the so-called **collapse** of the wave function upon measurement, introduced in 1927 by Heisenberg [117]. Collapse to what is controversial. The authoritative 2007 book by Schlosshauer [265] takes the "jump into an eigenstate", postulated by Dirac in the above quote, to be part of what he calls the "standard interpretation" of quantum mechanics. On the other hand, the older but also authoritative 1977 book by Landau & Lifschitz [171] explicitly remarks in the discussion in Section 7 that the state after the measurement is in general not an eigenstate.

The collapse requirement contradicts the unitary evolution of pure states through the Schrödinger equation (which is the mode of change of a closed system, hence in

the absence of a measurement), and depends on a not further detailed notion of measurement residing on the classical side of the cut. What happens to the state while the experiment is in progress but not complete is not specified. This makes the Copenhagen interpretation an incomplete description of the full temporal behavior of the state.

This incompleteness is a sign that we actually deal with a coarse-grained, reduced description. In such a reduced description, the description of the state of a particle is different before and after it passes a filter (polarizer, magnet, double slit, et cetera). The new information that the particle passed the filter requires a different description analogous to that responsible for the use of classical conditional probability when additional information arrives. In the quantum case, this is modeled by the collapse of the wave function. Landau's general case is the one modeled by event-based filters (Section 11.5), whereas Dirac's situation is modeled by the special case, where the filter operators R_k are the spectral projectors of an ideal measurement (Section 11.8).

In the thermal interpretation, collapse results from coarse-graining when the latter produces a reduced stochastic description in the form of a PDP (see Section 11.7). For example, we had seen in Section 7.8 that quantum-classical models may result from coarse-graining, and Bonilla & Guinea [42] give an explicit quantum-classical model that exhibits chaos and collapse.

Understanding that collapse comes from coarse-graining is a similar insight as that friction comes from coarse graining, an insight familiar to classical mechanics treated in the Markov approximation for a few relevant quantities. In both cases, the insight bridges the difference in the dynamics of an isolated system and that on an open system obtained by hiding the environment, turning it into a source of stochastic events. The explanation by coarse graining is in both cases fully quantitative and consistent with experiment, hence has all the features a good scientific explanation should have.

16.3 The minimal statistical interpretation

I reject the basic idea of contemporary statistical quantum theory, insofar as I do not believe that this fundamental concept will provide a useful basis for the whole of physics. [...]
One arrives at very implausible theoretical conceptions, if one attempts to maintain the thesis that the statistical quantum theory is in principle capable of producing a complete description of an individual physical system. On the other hand, those difficulties of theoretical interpretation disappear, if one views the quantum-mechanical description as the description of ensembles of systems. [...]
Within the framework of statistical quantum theory there is no such thing as a complete description of the individual system. [...]
it appears unavoidable to look elsewhere for a complete description of the individual system. [...]
I am rather firmly convinced that the development of theoretical physics will be of this type; but the path will be lengthy and difficult. [...]
the expectation that the adequate formulation of the universal laws involves the use of all conceptual elements which are necessary for a complete description, is more natural. [...]

If it should be possible to move forward to a complete description, it is likely that the laws would represent relations among all the conceptual elements of this description which, per se, have nothing to do with statistics.

Albert Einstein, 1949 [80, p. 673]

The Statistical Interpretation, according to which a pure state (and hence also a general state) provides a description of certain statistical properties of an ensemble of similarily prepared systems. [...]
In general, quantum theory predicts nothing which is relevant to a single measurement (excluding strict conservation laws like those of charge, energy, or momentum), and the result of a calculation pertains directly to an ensemble of similar measurements. For example, a single scattering experiment consists in shooting a single particle at a target and measuring its angle of scatter. Quantum theory does not deal with such an experiment, but rather with the statistical distribution (the differential cross section) of the results of an ensemble of similar experiments.

Leslie Ballentine, 1970 [22, pp. 360f]

We can now define the scope of quantum theory: In a strict sense, quantum theory is a set of rules allowing the computation of probabilities for the outcomes of tests which follow specified preparations. [...]
The above strict definition of quantum theory (a set of rules for computing the probabilities of macroscopic events) is not the way it is understood by most practicing physicists. They would rather say that quantum theory is used to compute the properties of microscopic objects, for example the energy-levels and cross-sections of atoms and nuclei. The theory can also explain some properties of bulk matter, such as the specific heat of solids or the electric conductivity of metals – whenever these macroscopic properties can be derived from those of the microscopic constituents. Despite this uncontested success, the epistemological meaning of quantum theory is fraught with controversy, perhaps because it is formulated in a language where familiar words are given unfamiliar meanings. Do these microscopic objects – electrons, photons, etc. – really exist, or are they only a convenient fiction introduced to help our reasoning, by supplying intuitive models in circumstances where ordinary intuition is useless?

Asher Peres, 2002 [233, p. 13]

That quantum mechanical states should be interpreted statistically goes back to 1926. Born [43, 44] had shown how the known statistical properties of scattering events are, in some sense, consistent with the deterministic Schrödinger equation, and can be derived from it assuming a statistical interpretation of the wave function. This earned him later a Nobel prize. But in Born's view, particles existed always in joint eigenstates of Hamiltonian and momentum that were modified discontinuously by random **quantum jumps**. In this way, the exact validity of the conservation laws of energy and momentum could be asserted.

Statistical interpretations in the precise sense (S) defined at the beginning of Chapter 16, have their beginnings with Weyl [300] and were discussed extensively by Ballentine [22], who contrasted it to the Copenhagen interpretation. In these statistical interpretations, a single (few or many particle) system has no state. Instead, the state is a property of the ensemble; one only talks about the prepared and observed properties of a population of experiments making up the ensemble. Equivalently,

the preparation procedure (which defines the ensemble on which measurements are performed) has, or defines, a state.

The **minimal interpretation** is a statistical interpretation (S) augmented by the additional stipulation that quantum mechanics is completely silent about a single system, and hence says nothing at all about it.[4]

According to Ballentine [22, pp. 366], for a consistent statistical interpretation, the notion of **preparation** must be clearly distinguished from that of **measurement**: *"State preparation refers to any procedure which will yield a statistically reproducible ensemble of systems. The concept of state in quantum theory can be considered operationally as an abbreviation for a description of the state preparation procedure. Of course there may be more than one experimental procedure which yields the same statistical ensemble, i. e., the same state. An important special case (which is sometimes incorrectly identified with measurement) is a filtering operation, which ensures that if a system passes through the filter it must immediately afterward have a value of some particular observable within a restricted range of its eigenvalue spectrum. On the other hand, measurement of some quantity E for an individual system means an interaction between the system and a suitable apparatus, so that we may infer the value of R (within some finite limits of accuracy) which the system had immediately before the interaction (or the value of R which the system would have had if it had not interacted, allowing for the possibility that the interaction will disturb the system)."* The **filtering** mentioned replaces the collapse in the Copenhagen interpretation, and serves the same purpose. In the thermal interpretation, it is modeled by event-based filters (Section 11.5).

The thermal interpretation of the situation described is that the preparation defines a state of the quantum fields present in the description. Their interaction with the detector produces observable events, whose statistics measures properties of the fields. In principle, quantum tomography (see Section 11.5) can be used to calibrate sufficiently stationary unknown sources, so that one can be sure which state they prepare in which setting. Knowing what was prepared and how to control it systematically, one can collect event statistics for new experimental settings and establish—on

4 True minimality is rare. Einstein [80] finds only the minimal interpretation consistent, but takes this as a limitation of quantum mechanics and expects the existence of a deeper underlying deterministic description. Ballentine [22] is not minimal throughout (despite an attempt to be so), as he assumes (p. 361) that definite positions exist: *"The Statistical Interpretation considers a particle to always be at some position in space, each position being realized with relative frequency $|\psi(r)|^2$ in an ensemble of similarily prepared experiments."* Even the most consequent book by Peres makes an exception at the very end (pp. 424f): *"This would cause no conceptual difficulty with quantum theory if the Moon, the planets, the interstellar atoms, etc., had a well defined state ρ. However, I have insisted throughout this book that ρ is not a property of an individual system, but represents the procedure for preparing an ensemble of such systems. How shall we describe situations that have no preparer? [...] Thus, a macroscopic object effectively [...] mimics, with a good approximation, a statistical ensemble. [...] You must have noted the difference between the present pragmatic approach and the dogmas held in the early chapters of this book."*

the basis of the resulting experimental evidence—a relation (10.2) between measurement results and properties of the system measured.

The single systems that allegedly travel from the source to the detector (but according to the minimal interpretation without any quantum properties) never enter the description, hence one cannot (and need not) say anything about these.

In a statistical interpretation, all statements claimed about single quantum systems are nonminimal. In particular, unlike the thermal interpretation, the minimal interpretation does not address the foundational problems posed by the ensembles of equilibrium thermodynamics (see Chapter II [204, Section 2.4]). Indeed, Ballentine [22, pp. 361] writes: *"The ensembles contemplated here are different in principle from those used in statistical thermodynamics, where we employ a representative ensemble for calculations, but the result of a calculation may be compared with a measurement on a single system. [...] Because the ensemble in the statistical interpretation is not merely a representative or calculational device, but rather it can and must be realized experimentally, it does not inject into quantum theory the same conceptual problems posed in statistical thermodynamics."*

Bibliography

In the bibliography, the number(s) after each reference give the page number(s) where it is cited.

[1] J. Agler and J. E. McCarthy, Complete Nevanlinna–Pick kernels, J. functional analysis, 175 (2000), 111–24. ⟨58⟩

[2] J. Agler, F. B. Yeh and N. J. Young, Reproducing kernel spaces and applications. Birkhäuser, Basel 2003. ⟨90⟩

[3] Y. Aharonov and D. Bohm, Significance of electromagnetic potentials in quantum theory, Phys. Rev. 115 (1959), 485–91. ⟨93⟩

[4] S. Akashi, Reproducing kernels and their applications. Springer, Berlin 1999. ⟨90⟩

[5] S. T. Ali, J.-P. Antoine and J.-P. Gazeau, Coherent states, wavelets and their generalizations, 2nd. ed., Springer, New York 2014. ⟨90⟩

[6] A. E. Allahverdyan, R. Balian and T. M. Nieuwenhuizen, Understanding quantum measurement from the solution of dynamical models, Physics Reports 525 (2013), 1–166. ⟨195, 235⟩

[7] A. E. Allahverdyan, R. Balian and T. M. Nieuwenhuizen, A sub-ensemble theory of ideal quantum measurement processes, Annals of Physics 376 (2017), 324–52. ⟨X, XIII, 32, 174, 195⟩

[8] D. Alpay, ed., Operator theory, Vol. 1. Springer, Basel 2015. ⟨90⟩

[9] D. Alpay, A. Dijksma, J. Rovnyak and H. de Snoo, Schur functions, operator colligations, and reproducing kernel Pontryagin spaces. Birkhäuser, Basel 1997. ⟨70, 90⟩

[10] J.-P. Antoine, S. T. Ali, W. Lisiecki, I. M. Mladenov and A. Odzijewicz (eds.), Quantization, coherent states, and complex structures. Springer, 1995. ⟨90⟩

[11] J. P. Antoine and C. Trapani, Partial inner product spaces: Theory and applications. Lecture notes in mathematics 1986, Springer, Heidelberg 2009. ⟨52⟩

[12] T. Appelquist and J. Carazzone, Infrared singularities and massive fields, Phys. Rev. D 11 (1975), 2856–61. ⟨237⟩

[13] H. Araki and M. M. Yanase, Measurement of quantum mechanical operators, Phys. Rev. 120 (1960), 622–6. ⟨237⟩

[14] N. Arcozzi, R. Rochberg, E. Sawyer and B. D. Wick, Distance functions for reproducing kernel Hilbert spaces, Function spaces in modern analysis. Contemp. Math 547 (2011), 25–53. ⟨83⟩

[15] P. N. Aronszajn, La théorie des noyaux reproduisants et ses applications, Premiére Partie, Math. Proc. Cambridge Phil. Soc. 39 (1943), 133–53. ⟨59, 70⟩

[16] N. Aronszajn, Theory of reproducing kernels, Trans. Amer. Math. Soc. 68 (1950), 337–404. ⟨59, 70⟩

[17] A. Aspect, Proposed experiment to test the nonseparability of quantum mechanics, Phys. Rev. D 14 (1976), 1944–1951. (Reprinted in [301].). ⟨159, 223⟩

[18] J. C. Baez, I. E. Segal and Z. Zhou, Introduction to algebraic and constructive quantum field theory, Princeton Univ. Press 1992, 2014. ⟨72⟩

[19] R. Balian, Incomplete descriptions and relevant entropies, Amer. J. Phys. 67 (1999), 1078–90. ⟨128⟩

[20] R. Balian, From microphysics to macrophysics: Methods and applications of statistical physics, 2 Vols, Springer, Berlin 2007. ⟨124, 127⟩

[21] J. A. Ball and V. Bolotnikov, De Branges–Rovnyak spaces: basics and theory, pp. 631–79 in: Operator theory (D. Alapay, ed.), Springer, Basel 2015. ⟨83⟩

[22] L. E. Ballentine, The statistical interpretation of quantum mechanics, Rev. Mod. Phys. 42 (1970), 358–81. ⟨206, 254, 255, 256⟩

[23] L. E. Ballentine, Quantum mechanics – a modern development, World Scientific, Singapore

https://doi.org/10.1515/9783110667387-017

1998. ⟨206⟩
[24] V. Bargmann, On a Hilbert space of analytic functions and an associated integral transform, Part I, Comm. Pure Appl. Math. 14 (1961), 187–214. ⟨72⟩
[25] D. Bar-Moshe and M. S. Marinov, Berezin quantization and unitary representations of Lie groups, Manuscript (1994). http://arxiv.org/abs/hep-th/9407093. ⟨102⟩
[26] R. J. Bartlett and M. Musial, Coupled-cluster theory in quantum chemistry, Rev. Mod. Phys. 79 (2007), 291–352. ⟨108⟩
[27] A. O. Barut and R. Raczka, Theory of group representations and applications, 2nd. ed., Warzawa 1980. ⟨98⟩
[28] E. Bauer, Théorie quantique de la valence. Les liaisons homéopolaires, Bull. Soc. Chim. France. Mém. (5) 1 (1934), 293–347. (Cited after [317]). ⟨230⟩
[29] M. B. Bekka and P. de la Harpe, Irreducibility of unitary group representations and reproducing kernels Hilbert spaces, Expositiones Math. 21 (2003), 115–49. ⟨90⟩
[30] J. S. Bell, On the Einstein Podolsky Rosen paradox, Physics 1 (1964), 195–200. (Reprinted in [301].). ⟨159⟩
[31] J. S. Bell, Subject and object, pp. 687–90 in: The physicist's conception of Nature (J. Mehra, ed.), Reidel, Dordrecht 1973. Reprinted pp. 40–44 in: Speakable ans unspeakable in quantum mechanics, Cambridge University Press, Cambridge 1987. ⟨144, 145⟩
[32] J. Bell, Against measurement, Physics World 3 (1990), 33–40. ⟨136⟩
[33] F. A. Berezin, Quantization in complex symmetric spaces, Math. USSR Izvestiya 9 (1975), 341–79. ⟨68⟩
[34] C. Berg, J. P. R. Christensen and P. Ressel, Harmonic analysis on semigroups. Springer, New York 1984. ⟨56, 70, 90⟩
[35] S. Bergman, Über die Kernfunktionen eines Bereiches und ihr Verhalten am Rande, J. reine angew. Math. 169 (1933), 1–42. ⟨59⟩
[36] A. N. Beris and B. J. Edwards, Thermodynamics of flowing systems with internal microstructure, Oxford Univ. Press, New York 1994. ⟨128⟩
[37] A. Berlinet and C. Thomas-Agnan, Reproducing kernel Hilbert spaces in probability and statistics. Springer, 2004. ⟨90⟩
[38] S. Bochner, Über orthogonale Systeme analytischer Funktionen, Math. Zeitschrift 14 (1922), 180–207. ⟨62⟩
[39] N. Bohr, The quantum postulate and the recent development of atomic theory, Lecture delivered on Sept. 16, 1927 in Como. Published in: Nature April 14 (1928), 580–90. ⟨158, 210, 249, 250⟩
[40] N. Bohr, On the notions of causality and complementarity, Dialectica 2 (1948), 312. Reprinted in: Science, New Ser. 111 (1950), 51–54. ⟨222⟩
[41] L. Boltzmann, Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten, Wiener Berichte 58 (1868), 517–60. ⟨42⟩
[42] L. L. Bonilla and F. Guinea, Collapse of the wave packet and chaos in a model with classical and quantum degrees of freedom, Phys. Rev. A 45 (1992), 7718–28. ⟨253⟩
[43] M. Born, Zur Quantenmechanik der Stoßvorgänge. Zeitschrift f. Physik 37 (1926), 863–7. (English translation: On the quantum mechanics of collisions, Section I.2, pp. 52–55 in [301].). ⟨230, 231, 254⟩
[44] M. Born, Quantenmechanik der Stoßvorgänge, Z. Phys. 38 (1926), 803–27. ⟨230, 231, 247, 254⟩
[45] M. Born, Das Adiabatenprinzip in der Quantenmechanik, Z. Phys. 40 (1927), 167–92. ⟨135, 230, 231⟩
[46] M. Born, Über den Sinn der physikalischen Theorien, Naturwissenschaften 17 (1929), 109–18. ⟨251⟩

[47] M. Born and W. Heisenberg, La mécanique des quanta, pp. 143–78 in: Électrons et photons. Rapports et discussions du cinquième Conseil de physique tenu à Bruxelles du 24 au 29 Octobre 1927 sous les auspices de l'Institut international de physique Solvay (H. A. Lorentz, ed.), Gauthier-Villars, Paris 1928. ⟨159⟩
[48] S. N. Bose, Plancks Gesetz und Lichtquantenhypothese, Zeitschrift f. Physik 26 (1924), 178–81. ⟨207⟩
[49] W. Boucher and J. Traschen, Semiclassical physics and quantum fluctuations, Phys Rev. D 37 (1988), 3522–32. ⟨160⟩
[50] L. de Branges, Hilbert spaces of entire functions. Prentice-Hall, 1968. ⟨82, 90⟩
[51] L. de Branges and J. Rovnyak, Canonical models in quantum scattering theory, pp. 295–392 in: Perturbation theory and its applications in quantum mechanics (C. H. Wilcox, ed.), Wiley, New York 1966. ⟨82⟩
[52] L. de Branges and J. Rovnyak, Square summable power series. Courier Corporation, 2015. ⟨83⟩
[53] H. Breuer and F. Petruccione, Quantum measurement and the transformation from quantum to classical probabilities, Phys. Rev. A 54 (1996), 1146–53. ⟨129⟩
[54] H.-P. Breuer and F. Petruccione, The theory of open quantum systems. Oxford Univ. Press, New York, 2002. ⟨163, 174, 196⟩
[55] A. E. Brouwer, A. M. Cohen, and A. Neumaier, Distance-regular graphs, Ergebnisse der Mathematik und ihrer Grenzgebiete, 3rd ed., vol. 18, Springer, Berlin 1989. ⟨10, 89, 90⟩
[56] D. Buchholz, The physical state space of quantum electrodynamics, Comm. Math. Phys. 85 (1982), 49–71. ⟨237⟩
[57] Bureau International des Poids et Mesure, The international system of units (SI), 9th edition 2019, Draft, 18 December 2019. https://www.bipm.org/utils/en/pdf/si-revised-brochure/Draft-SI-Brochure-2018.pdf. ⟨176, 221⟩
[58] H. B. Callen. Thermodynamics and an introduction to thermostatistics, 2nd ed., Wiley, New York, 1985. (The quote from Callen can also be found at the end of Section 2 of the article http://www.polyphys.mat.ethz.ch/education/lec_thermo/callen_article.pdf.). ⟨104, 141, 142, 155⟩
[59] E. Calzetta and B. L. Hu, Nonequilibrium quantum field theory, Cambridge Univ. Press, New York 2008. ⟨114, 124, 155, 198, 201⟩
[60] M. Combescure and D. Robert, Coherent states and applications in mathematical physics, Springer, 2012. ⟨90⟩
[61] A. Connes, Une classification des facteurs de type III, Ann. Sci. Ecole Norm. Sup. 6 (1973), 133–252. ⟨245⟩
[62] R. Courant and D. Hilbert, Methoden der mathematischen Physik, Springer, Berlin, Band I, 1924, Band II, 1930. ⟨138⟩
[63] Currie, Jordan and Sudarshan, Relativistic invariance and Hamiltonian theories of interacting particles, Rev. Mod. Phys. 35 (1963), 350–75. ⟨213⟩
[64] K. K. Darrow, Elementare Einführung in die Wellenmechanik, Hirzel, Leipzig 1929. ⟨247⟩
[65] E. B. Davies and H. Spohn Open quantum systems with time-dependent Hamiltonians and their linear response, J. Statistical Physics 19 (1978), 511–23. ⟨120⟩
[66] G. Dell'Antonio, R. Figari and A. Teta, A time-dependent perturbative analysis for a quantum particle in a cloud chamber, Ann. Henri Poincaré 11 (2010), 539–64. ⟨205⟩
[67] A. Der Kiureghian and O. Ditlevsen, Aleatory or epistemic? Does it matter? Structural Safety 31 (2009), 105–12. ⟨33⟩
[68] P. A. M. Dirac, The fundamental equations of quantum mechanics, Proc. Roy. Soc. Lond. A 109 (1925), 642–53. ⟨19⟩
[69] P. A. M. Dirac, Note on exchange phenomena in the Thomas atom, Proc. Cambridge Philos.

Soc. 26 (1930), 376–85. ⟨93, 124⟩
[70] P. A. M. Dirac, The principles of quantum mechanics, 1st ed., Oxford Univ. Press, Oxford 1930. ⟨X, 26, 233, 235, 252⟩
[71] P. A. M. Dirac, The principles of quantum mechanics, 3rd ed., Clarendon Press, Oxford 1947. ⟨233⟩
[72] P. A. M. Dirac, Development of the physicist's conception of Nature, Section 1, pp. 1–14 in: The physicist's conception of Nature (J. Mehra, ed.), Reidel, Dordrecht 1973. ⟨144, 247⟩
[73] P. A. M. Dirac, The development of quantum mechanics, pp. 1–20 in: Directions in Physics (H. Hora and J. R. Shepanski, eds.), Wiley, New York 1978. ⟨154⟩
[74] P. D. Drummond and D. F. Walls, Quantum theory of optical bistability I. Nonlinear polarisability model. J. Physics A Math. Gen. 13 (1980), 725–41. ⟨199⟩
[75] P. H. Eberhard and R. R. Ross, Quantum field theory cannot provide faster-than-light communication, Foundations of Physics Letters 2 (1989), 127–49. ⟨223⟩
[76] P. Ehrenfest, Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik, Zeitschrift für Physik 45 (1927), 455–7. ⟨9, 21, 23, 24, 151⟩
[77] A. Einstein, Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt, Annalen der Physik 322 (1905), 132–48. ⟨207⟩
[78] A. Einstein, Maxwell's influence on the development of the conception of physical reality, pp. 66–73 in: James Clerk Maxwell: A commemorative volume 1831–931, Macmillan, New York, 1931. http://www.d-meeus.be/physique/Maxwell-Einstein-en.html. ⟨201⟩
[79] A. Einstein, On the method of theoretical physics, Philosophy of Science 1 (1934), 163–9. ⟨137, 151, 211⟩
[80] A. Einstein, Reply to criticisms, pp. 665–88 in: Albert Einstein: philosopher – scientist (P. A. Schilpp, ed.), Cambridge Univ. Press, Cambridge 1949. ⟨135, 144, 157, 247, 254, 255⟩
[81] A. Einstein, Elementare Überlegungen zur Interpretation der Grundlagen der Quanten-Mechanik, pp. 33–40 in: Scientific papers presented to Max Born, Oliver & Boyd, Edinburgh 1953. ⟨154⟩
[82] M. Engliš, Berezin quantization and reproducing kernels on complex domains, Trans. Amer. Math. Soc. 348 (1996), 411–79. ⟨102⟩
[83] M. Engliš, Quantization methods: A guide for physicists and analysts, Rev. Math. Phys. 17 (2005), 391–490. math-ph/0405065. ⟨102⟩
[84] R. C. Falcao and F. Parisio, Fragmentation of brittle plates by localized impact, Applied Physics Letters 105 (2014), 124102. ⟨205⟩
[85] J. Faraut and A. Korányi, Analysis on symmetric cones. Clarendon Press, Oxford 1994. ⟨59, 68, 89, 90⟩
[86] J. C. Ferreira and V. A. Menegatto, Eigenvalues of integral operators defined by smooth positive definite kernels, Integral equation and Operator Theory, 64 (2009), 61–81. ⟨62⟩
[87] R. Figari and A. Teta, Emergence of classical trajectories in quantum systems: the cloud chamber problem in the analysis of Mott (1929), Arch. Hist. Exact Sci. 67 (2013), 215–34. ⟨205⟩
[88] C. H. FitzGerald and R. A. Horn, On fractional Hadamard powers of positive definite matrices, J. Math. Anal. Appl. 61 (1977), 633–42. ⟨68⟩
[89] V. Fock, Konfigurationsraum und zweite Quantelung, Zeitschrift für Physik 75 (1932), 622–47. ⟨72⟩
[90] G. B. Folland, Real analysis. Modern techniques and their applications. John Wiley, New York 1999. ⟨61⟩
[91] I. I. Frenkel, Wave mechanics, advanced general theory, Clarendon Press, Oxford 1934, pp. 253, 436. ⟨93, 124⟩
[92] J. Fröhlich, A brief review of the "ETH-approach to quantum mechanics", Manuscript (2019).

https://arxiv.org/pdf/1905.06603.pdf. ⟨6, 247⟩
[93] J.-P. Gazeau, Coherent states in quantum physics. Wiley-VCH, Weinheim 2009. ⟨90⟩
[94] I. M. Gelfand and M. A. Naimark, On the imbedding of normed rings into the ring of operators in Hilbert spaces (in Russian). Mat. Sbornik 12 (1943), 197–213. ⟨45⟩
[95] P. Gérard, P. A. Markowich, N. J. Mauser and F. Poupaud, Homogenization limits and Wigner transforms, Comm. Pure Appl. Math. 50 (1997), 323–79. ⟨131⟩
[96] J. W. Gibbs, Elementary principles in statistical mechanics, Yale Univ. Press, New Haven, 1902. (Citations from the Reprinted, Dover 1960). ⟨42, 243, 244⟩
[97] J. Ginibre, On the asymptotic exactness of the Bogoliubov approximation for many Boson systems, Comm. Math. Phys. 8 (1968), 26–51. ⟨25⟩
[98] R. Glauber, Coherent and incoherent states of the radiation field, Phys. Rev. 131 (1963), 2766–88. Reprinted as pp. 202–224 in [161]. ⟨9, 69⟩
[99] J. Glimm and A. Jaffe, Quantum physics – a functional integral point of view, 2nd ed., Springer, New York 1987. ⟨72⟩
[100] T. Gneiting and A. E. Raftery, Weather forecasting with ensemble methods, Science 310 (2005), 248–9. ⟨42⟩
[101] C. Godsil, Algebraic combinatorics. Chapman and Hall, New York 1993. ⟨90⟩
[102] A. L. Goodman, Hartree–Fock–Bogoliubov theory with applications to nuclei, Adv. Nucl. Phys. (United States) 11 (1979), 263–366. ⟨108⟩
[103] K. Gottfried, Does quantum mechanics carry the seeds of its own destruction? Physics World 4 (1991), 34–40. ⟨3, 235⟩
[104] H. Grabert, Projection operator techniques in nonequilibrium statistical mechanics, Springer Tracts in Modern Physics, Berlin 1982. ⟨124⟩
[105] J. M. Gracia-Bondía, J. C. Várilly, Quantum electrodynamics in external fields from the spin representation, J. Math. Phys. 35, 1994, 3340–67. ⟨117⟩
[106] C. W. J. Granger and P. Newbold, Forecasting economic time series, Academic Press, New York 1977. ⟨42⟩
[107] F. Grassi, Particle emission in hydrodynamics: a problem needing a solution, Brazilian J. Physics 35 (2005), 52–9. ⟨252⟩
[108] J. J. Griffin and J. A. Wheeler, Collective motions in nuclei by the method of generator coordinates, Phys. Rev. 108 (1957), 311–27. ⟨108⟩
[109] D. J. Griffiths and D. F. Schroeter, Introduction to quantum mechanics, Cambridge University Press, Cambridge 2018. ⟨179⟩
[110] R. Haag, Quantum field theories with opposite particles and asymptotic conditions, Phys. Rev. 112 (1958), 669–73. ⟨122⟩
[111] F. Haake, Quantum signatures of chaos, 3rd ed., Springer, Heidelberg 2010. ⟨194⟩
[112] P. Hänggi, P. Talkner and M. Borkovec, Reaction-rate theory: fifty years after Kramers, Rev. Modern Phys. 62 (1990), 251–341. ⟨199⟩
[113] P. Hänggi and H. Thomas, Stochastic processes: Time evolution, symmetries and linear response, Physics Reports, 88 (1982), 207–319. ⟨120⟩
[114] M. J. W. Hall, D. Deckert and H. M. Wiseman, Quantum phenomena modeled by interactions between many classical worlds, Phys. Review X4 (2014), 041013. ⟨151⟩
[115] D. J. Heeger and J. R.Bergen, Pyramid-based texture analysis/synthesis, pp. 229–38 in: Proc. 22nd Annual ACM Conf. Computer Graphics and Interactive Techniques, ACM 1995. ⟨42⟩
[116] G. C. Hegerfeldt, Instantaneous spreading and Einstein causality in quantum theory, Annalen der Physik 7 (1998), 716–25. ⟨237⟩
[117] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, Z. Phys. 43 (1927), 172–98. (English translation: The physical content of quantum kinematics and mechanics, Section I.3, pp. 62–84 in [301].). ⟨33, 157, 206, 207, 250, 251,

252⟩
[118] W. Heisenberg, Discussione sulla communicazione Bohr, pp. 589–98 in: Proc. Int. Conf. Physicists, Como 1927. ⟨249, 250⟩
[119] W. Heisenberg, Die Entwicklung der Quantentheorie 1918–1928, Naturwissenschaften 17 (1929), 490–6. ⟨165⟩
[120] W. Heisenberg, Die physikalischen Prinzipien der Quantentheorie, Hirzel, Leipzig 1930. (English translation: The physical principles of the quantum theory, Univ. of Chicago Press, Chicago 1930.). ⟨235⟩
[121] W. Heisenberg, Wandlungen der Grundlagen der exakten Naturwissenschaft in jüngster Zeit, Naturwissenschaften 22 (1934), 669–75. ⟨249⟩
[122] W. Heisenberg, Der Teil und das Ganze, Piper, München 1969. (English translation: W. Heisenberg, Physics and beyond: encounters and conversations, Harper & Row 1972.). ⟨153⟩
[123] C. W. Helstrom, Quantum detection and estimation theory, Academic Press, New York 1976. ⟨192⟩
[124] A. Herdegen, Semidirect product of CCR and CAR algebras and asymptotic states in quantum electrodynamics, J. Math. Phys. 39 (1998), 1788–817. ⟨117⟩
[125] A. Herdegen, Asymptotic structure of electrodynamics revisited, Lett. Math. Phys. 107 (2017), 1439–70. ⟨117⟩
[126] P. Hertz, Über die mechanischen Grundlagen der thermodynamik, Ann. Physik IV. Folge (33) 1910, 225–74. ⟨243⟩
[127] C. S. Herz, Une ébauche d'une théorie générale des fonctions définies-négatives, Séminaire Brelot–Choquet–Deny. Théorie du potentiel 7(3) (1962–1963), 1–10. ⟨66⟩
[128] T. Hida and S. Si, Lectures on white noise functionals, World Scientific, Singapore 2008. ⟨72, 90⟩
[129] T. Hida and S. Streit, Let us use white noise functionals, World Scientific, Singapore 2017. ⟨72, 90⟩
[130] D. G. Higman, Coherent configurations. Part I: Ordinary representation theory, Geometriae Dedicata 4 (1975), 1–32. ⟨10⟩
[131] D. G. Higman, Coherent algebras, Linear Algebra Appl. 93 (1987), 209–39. ⟨10⟩
[132] D. Hilbert, Mathematische Probleme, Nachrichten Ges. Wiss. Göttingen, Math.-Phys. Klasse, 1900 (1900), 253–97. English transl.: Mathematical problems, Bulletin AMS 8 (1902), 437–479. ⟨137, 138⟩
[133] D. Hilbert, J. v. Neumann and L. Nordheim, Über die Grundlagen der Quantenmechanik, Math. Annalen 98 (1928), 1–30. ⟨138⟩
[134] A. Hobson, There are no particles, there are only fields. Amer. J. Physics 81 (2013), 211–23. ⟨208⟩
[135] R. Hoffmann, V. I. Minkin and B. K. Carpenter, Ockham's razor and chemistry. HYLE Int. J. Phil. Chem 3 (1997), 3–28. ⟨159⟩
[136] M. O. Hongler and W. M. Zheng, Exact solution for the diffusion in bistable potentials, J. Statist. Phys, 29 (1982), 317–27. ⟨179⟩
[137] G. 't Hooft, Quantum gravity as a dissipative deterministic system, Class. Quantum Gravity 16 (1999), 3263–79. ⟨247⟩
[138] E. Hopf, Über die Bedeutung der willkürlichen Funktionen für die Wahrscheinlichkeitstheorie, Jahresbericht DMV 56 (1936), 179–95. ⟨42⟩
[139] R. A. Horn, The theory of infinitely divisible matrices and kernels, Trans. Amer.Math. Soc. 136 (1969), 269–86. ⟨66⟩
[140] A. Horzela and F. H. Szafraniec, A measure-free approach to coherent states, J. Phys. A: Math. Theor. 45 (2012), 244018. ⟨90⟩
[141] L. van Hove, Sur certaines représentations unitaires d'un groupe infini de transformations (in

French). Proc. Roy. Acad. Sci. Belgium 26 (1951), 1–102. ⟨102⟩
[142] F. Iachello, Lie algebras and applications, Springer, Berlin 2006. ⟨100⟩
[143] P. Iglesias-Zemmour, Diffeology, Mathematical Surveys and Monographs Vol. 185, American Mathematical Soc. 2013. ⟨77⟩
[144] A. Inomata, H. Kuratsuji and C. C. Gerry, Path integrals and coherent states of SU(2) and SU(1, 1). World Scientific, 1992. ⟨90⟩
[145] R. Jackiw, Quantum meaning of classical field theory, Rev. Mod. Phys. 49, 681–706 (1977). ⟨118⟩
[146] E. T. Jaynes, Information theory and statistical mechanics, Phys. Rev. 106 (1957), 620–30. ⟨158⟩
[147] J. H. Jeans, Report on radiation and the quantum theory, 2nd ed., Fleetway, London 1924. ⟨203⟩
[148] S. Jeon and U. Heinz, Introduction to hydrodynamics, Int. J. Modern Physics E 24 (2015), 1530010. arXiv:1503.03931. ⟨252⟩
[149] S. Jeon and L. G. Yaffe, From quantum field theory to hydrodynamics: Transport coefficients and effective kinetic theory, Phys. Review D 53 (1996), 5799–809. ⟨123, 250⟩
[150] E. Joos and H. D. Zeh, The emergence of classical properties through interaction with the environment, Zeitschrift f. Physik B Condensed Matter 59 (1985), 223–43. ⟨174⟩
[151] P. Jordan, Über eine neue Begründung der Quantenmechanik, Z. Phys. 40 (1927), 809–38. ⟨230, 231⟩
[152] D. Kapec, M. Perry, A. Raclariu and A. Strominger, Infrared divergences in QED, revisited, Phys. Rev. D 96 (2017), 085002. ⟨117⟩
[153] R. Kapral and G. Ciccotti, Mixed quantum-classical dynamics, J. Chem. Phys. 110 (1990), 8919–29. ⟨129⟩
[154] B. D. Keister and W. N. Polyzou, Relativistic Hamiltonian dynamics in nuclear and particle physics. Adv. Nuclear Physics 20 (1991), 226–479. ⟨237⟩
[155] G. A. Kerimov, New algebraic approach to scattering problems, Phys. Rev Lett. 80 (1998), 2976–9. ⟨112⟩
[156] G. A. Kerimov, Intertwining operators and S matrix, Phys. Atomic Nuclei 65 (2002), 1036–41. ⟨112⟩
[157] J. R. Klauder, Continuous-representation theory. II. Generalized relation between quantum and classical dynamics. J. Math. Phys. 4 (1963), 1058–73. ⟨71⟩
[158] J. R. Klauder, Continuous-representation Theory. III. On functional quantization of classical systems, J. Math. Phys. 5 (1964), 177–87. ⟨85, 94⟩
[159] J. R. Klauder, Are coherent states the natural language of quantum theory? pp. 1–12 in: Fundamental aspects of quantum theory (V. Gorini et al., eds.) Plenum, New York 1986. ⟨8⟩
[160] J. Klauder, Coherent states in action, Manuscript (1997). arXiv:quant-ph/9710029. ⟨VII, 156⟩
[161] J. R. Klauder and B.-S. Skagerstam (eds.), Coherent states. Applications in physics and mathematical physics, World Scientific, Singapore 1985. ⟨90, 261⟩
[162] W. H. Klink, Point form relativistic quantum mechanics and electromagnetic form factors, Phys. Rev. C 58, 3587–604 (1998). ⟨118⟩
[163] A. N. Kolmogorov, Stationary sequences in Hilbert space, Bull. Math. Univ. Moscow, 2 (1941), 1–40. ⟨59⟩
[164] A. Kolmogorov, Foundations of the theory of probability, 2nd English ed., Chelsea, New York, 1956. ⟨31⟩
[165] B. O. Koopman, Hamiltonian systems and transformations in Hilbert space, Proc. Nat. Acad. Sci. 17 (1931), 315–8. ⟨132⟩
[166] P. Kramer and M. Saraceno, Geometry of the time-dependent variational principle in quantum mechanics, Springer, Berlin 1981. ⟨93, 94, 124⟩

[167] M. G. Kreĭn, Hermitian-positive kernels on homogeneous spaces. I, Amer. Math. Soc. Transl. 2 (1963), 69–108. ⟨63, 70⟩
[168] M. G. Kreĭn, Hermitian-positive kernels on homogeneous spaces. II, Amer. Math. Soc. Transl. 2 (1963), 109–64. ⟨70⟩
[169] L. Krüger, G. Gigerenzer and M. S. Morgan (eds.), The probabilistic revolution: Ideas in the sciences, Vol. 2, MIT Press, Cambridge, MA, 1987. ⟨31⟩
[170] L. Landau, Das Dämpfungsproblem in der Wellenmechanik, Z. Phys. 45 (1927), 430–41. ⟨232⟩
[171] L. D. Landau and E. M. Lifshitz, Course of theoretical physics, Vol. 3: Quantum mechanics, 3rd ed., Pergamon Press, 1977. ⟨3, 234, 240, 252⟩
[172] L. D. Landau and E. M. Lifshitz, Course of theoretical physics, Vol. 5: Statistical mechanics, Part 1, 3rd ed., Pergamon Press, 1980. ⟨43, 240, 241, 242⟩
[173] E. H. Lieb and R. Seiringer, Justification of c-number substitutions in Bosonic Hamiltonians, Phys. Rev. Lett. 94 (2005), 080401. ⟨25⟩
[174] G. Lindblad, On the generators of quantum dynamical semigroups, Comm. Math. Phys. 48 (1976), 119–30. ⟨125, 127⟩
[175] W. Liu (ed.), Handbook of relativistic quantum chemistry, Springer, Berlin, 2017. ⟨208⟩
[176] E. N. Lorenz, On the existence of a slow manifold, J. Atmospheric Sci. 43 (1986), 1547–57. ⟨198⟩
[177] E.-L. Malus, Sur une propriété de la lumière réfléchie, Mém. Phys. Chim. Soc. D'Arcueil 2 (1809), 143–58. Reprinted and partially translated into English in [281]. ⟨148⟩
[178] L. Mandel and E. Wolf, Optical coherence and quantum optics, Cambridge Univ. Press, Cambridge 1995. ⟨150, 202, 203, 215, 216⟩
[179] H. Margenau, The nature of physical reality, Mc Graw-Hill, New York 1950. ⟨190⟩
[180] A. Martinez, An introduction to semiclassical and microlocal analysis, Springer, New York 2002. ⟨127⟩
[181] J. E. Marsden and T. S. Ratiu, Introduction to mechanics and symmetry, Springer, New York 1994. ⟨23, 99⟩
[182] K. Maurin, General eigenfunction expansions and unitary representations of topological groups, Polish Scientific Publishers, Warsaw 1968. ⟨25⟩
[183] A. D. McLachlan and M. A. Ball, Time-dependent Hartree–Fock theory for molecules, Rev. Mod. Phys. 36 (1964), 844–55. ⟨108⟩
[184] K. Menger, Untersuchungen über allgemeine Metrik, Math. Annalen 100 (1928), 75–163. ⟨65⟩
[185] J. Mercer, Functions of positive and negative type and their connection with the theory of integral equations, Philos. Trans. Roy. Soc. London (Ser. A) 209 (1909), 415–46. ⟨62, 70⟩
[186] E. H. Moore, General analysis, Vol. 1. Memoirs of the American Philosophical Society, 1935. ⟨59⟩
[187] N. F. Mott, The wave mechanics of α-ray tracks, Proc. Royal Soc. London A 126 (1929), 79–84. ⟨204, 205, 218⟩
[188] K.-H. Neeb, Holomorphy and convexity in Lie theory. de Gruyter, Berlin 2000. ⟨70, 89, 90⟩
[189] Y. Neretin, Lectures of Gaussian integral operators and classical groups. Europ. Math. Soc., 2011. ⟨70, 90⟩
[190] A. Neumaier, Distances, graphs, and designs. Europ. J. Combinatorics 1 (1980), 163–74. ⟨90⟩
[191] A. Neumaier, Classification of graphs by regularity, J. Combin. Theory (Ser. B) 30 (1981), 318–31. ⟨90⟩
[192] A. Neumaier, Distance matrices, dimension, and conference graphs, Proc. Kon. Ned. Akad. Wet., Ser. A, 84 = Indagationes Math. 43 (1981), 385–91. ⟨90⟩
[193] A. Neumaier, Derived eigenvalues of symmetric matrices, with applications to distance geometry, Linear Algebra Appl. 134 (1990), 107–20. ⟨90⟩
[194] A. Neumaier, Ensembles and experiments in classical and quantum physics, Int. J. Mod. Phys.

B 17 (2003), 2937–80. ⟨XII, 20, 238⟩
[195] A. Neumaier, Quantum field theory as eigenvalue problem, Unpublished manuscript (2003). http://arxiv.org/abs/gr-qc/0303037. ⟨20⟩
[196] A. Neumaier, The beginnings of the thermal interpretation, Web site. https://www.mat.univie.ac.at/~neum/physfaq/therm/beginnings.html. ⟨XIII⟩
[197] A. Neumaier, Optical models for quantum mechanics, Lecture given on February 16, 2010 at the Institut für Theoretische Physik, Universität Giessen, Germany. https://www.mat.univie.ac.at/~neum/ms/optslides.pdf. ⟨XIII⟩
[198] A. Neumaier, The thermal interpretation of quantum mechanics, Web site (since 2010). http://www.mat.univie.ac.at/~neum/physfaq/therm/thermalMain.html. ⟨XIII⟩
[199] A. Neumaier, The vacuum fluctuationmMyth, PhysicsForums Insights (November 2016). https://www.physicsforums.com/insights/vacuum-fluctuation-myth/. ⟨43, 210⟩
[200] A. Neumaier, Coherent spaces, Web site (2017). http://www.mat.univie.ac.at/~neum/cohSpaces.html. ⟨XI, 91⟩
[201] A. Neumaier, Review of "Curie-Weiss model of the quantum measurement process" by Allahverdyan, Balian, and Nieuwenhuizen, Posting on Physics Overflow, May 26, 2017. https://www.physicsoverflow.org/39105. ⟨195⟩
[202] A. Neumaier, Introduction to coherent spaces, Manuscript (2018). https://arxiv.org/abs/1804.01402. ⟨XI⟩
[203] A. Neumaier, Foundations of quantum mechanics I. A critique of the tradition, Manuscript (2019). https://arxiv.org/abs/1902.10778. ⟨XI⟩
[204] A. Neumaier, Foundations of quantum physics II. The thermal interpretation, Manuscript (2019). https://arxiv.org/abs/1902.10779. ⟨XI, 256⟩
[205] A. Neumaier, Foundations of quantum mechanics III. Measurement, Manuscript (2019). https://arxiv.org/abs/1902.10782. ⟨XI⟩
[206] A. Neumaier, Foundations of quantum physics IV. More on the thermal interpretation, Manuscript (2019). https://arxiv.org/abs/1904.12721. ⟨XI⟩
[207] A. Neumaier, Foundations of quantum physics V. Coherent foundations, Manuscript (2019). https://arxiv.org/abs/1905.00920. ⟨XI⟩
[208] A. Neumaier, Finite coherent spaces and finite geometries, Manuscript (2019). ⟨10⟩
[209] A. Neumaier, The classical limit in coherent spaces, Manuscript (2019). ⟨81⟩
[210] A. Neumaier, PhysicsForums posting (February 12, 2019). https://www.physicsforums.com/threads/against-interpretation-comments.965294/page-8#post-6132213. ⟨139⟩
[211] A. Neumaier, The 7 basic rules of quantum mechanics, PhysicsForums Insights (May 2019). https://www.physicsforums.com/insights/the-7-basic-rules-of-quantum-mechanics/. ⟨XI, 1⟩
[212] A. Neumaier and A. Ghaani Farashahi, Introduction to coherent quantization, Manuscript (2019). https://arxiv.org/abs/1804.01400. ⟨XI, 11, 71, 74, 75, 97, 101, 116⟩
[213] A. Neumaier and A. Ghaani Farashahi, Finite-dimensional coherent manifolds, Manuscript (2019). ⟨77⟩
[214] A. Neumaier and D. Westra, Classical and quantum mechanics via Lie algebras. Online book, 1st ed. 2008, 2nd ed. 2011. http://arxiv.org/abs/0810.1019. ⟨XIII, 7, 32, 34, 80, 99, 126, 132, 188⟩
[215] J. von Neumann, Mathematische Begründung der Quantenmechanik, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Klasse 1927 (1928), 1–57. ⟨232⟩
[216] J. von Neumann, Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Klasse 1927 (1928), 245–72. ⟨165, 232⟩
[217] J. von Neumann, Mathematische Grundlagen der Quantenmechanik, Springer, Berlin 1932. ⟨249, 250⟩
[218] M. A. Nielsen and I. L. Chuang Quantum computation and quantum information: 10th

Anniversary Edition, Cambridge Univ. Press, Cambridge 2011. ⟨179⟩
[219] M. M. Nieto, The discovery of squeezed states – in 1927, Manuscript (1997). https://arxiv.org/abs/1quant-ph/9708012. ⟨9⟩
[220] NIST, International and U.S. perspectives on measurement uncertainty, Web document. http://physics.nist.gov/cuu/Uncertainty/international1.html. ⟨31, 33, 235⟩
[221] NIST, Historical context of the SI, Web document. http://physics.nist.gov/cuu/Units/second.html. ⟨141⟩
[222] N. Obata, White noise calculus and Fock space, Lecture Notes in Mathematics Vol. 1577, Springer, Berlin 1994. ⟨72, 90⟩
[223] W. of Ockham, Summa totius logicae (ca. 1323) in: Philosophical writings (P. Boehner, ed.) Nelson, Edinburgh 1957. ⟨159⟩
[224] H. C. Oettinger, Beyond equilibrium thermodynamics, Wiley, Hoboken 2005. ⟨128⟩
[225] G. I. Ol'shanskii, Invariant cones in Lie algebras, Lie semigroups and the holomorphic discrete series, Funct. Anal. Appl. 15 (1981), 275–85. ⟨89⟩
[226] J. Ono and K. Ando, Semiquantal molecular dynamics simulations of hydrogen-bond dynamics in liquid water using multi-dimensional Gaussian wave packets, J. Chem. Phys. 137 (2012), 174503. ⟨108⟩
[227] M. Ozawa, Quantum measuring processes of continuous observables, J. Math. Phys. 25 (1984), 79–87. ⟨237⟩
[228] K. R. Parthasarathy and K. Schmidt, Positive definite kernels, continuous tensor products, and central limit theorems of probability theory. Lecture Notes in Mathematics 272, Springer, Berlin 1972. ⟨75, 90⟩
[229] M. E. Paté-Cornell, Uncertainties in risk analysis: Six levels of treatment, Reliability Engineering & System Safety 54 (1996), 95–111. ⟨33⟩
[230] A. K. Pattanayak and W. C. Schieve, Gaussian wave-packet dynamics: Semiquantal and semiclassical phase-space formalism, Phys. Rev. E 50 (1994), 3601–14. ⟨108⟩
[231] W. Pauli, Über Gasentartung und Paramagnetismus, Z. Physik 41 (1927), 81–102. ⟨231⟩
[232] A. M. Perelomov, Generalized coherent states and their applications, Springer, Berlin, 1986. ⟨70, 89, 90, 184, 192, 194⟩
[233] A. Peres, Quantum theory: Concepts and methods, Kluwer, Dordrecht 2002. ⟨179, 254⟩
[234] A. Peres and D. R. Terno, Hybrid classical-quantum dynamics, Phys. Rev. A 63 (2001), 022101. ⟨129⟩
[235] A. Peres and D. R. Terno, Quantum information and special relativity, Int. J. Quantum Information 1 (2003), 225–35. ⟨222, 225⟩
[236] A. Peres and D. R. Terno, Quantum information and relativity theory, Rev. Mod. Phys. 76 (2004), 93–123. ⟨225⟩
[237] J. J. G. Perez and R. Ossikovski, Polarized light and the Mueller matrix approach, CRC Press, Boca Raton, 2017. ⟨149⟩
[238] M. Planck, Entropy and temperature of radiant heat, Annalen der Physik 1 (1900), 719–37. ⟨26, 207⟩
[239] Plato, Timaeus. https://www.ellopos.net/elpenor/physis/plato-timaeus/genesis.asp?pg=2. ⟨137⟩
[240] J. von Plato, Creating modern probability, Cambridge Univ. Press, Cambridge 1994. ⟨31⟩
[241] H. Poincaré, La science et l'hypothèse, Flammarion, 1902. English translation: Science and hypothesis, Walter Scott, London 1905. ⟨161, 175, 178, 235⟩
[242] O. V. Prezdho, A quantum-classical bracket that satisfies the Jacobi identity, J. Chem. Phys. 124 (2006), 201104. ⟨129⟩
[243] O. V.Prezhdo and V. V. Kisil, Mixed quantum and classical mechanics, Phys. Rev. A 56 (1997), 162–75. ⟨129⟩

[244] A. A. Raduta, Nuclear structure with coherent states. Springer, 2015. ⟨90⟩
[245] J. Rau and B. Müller, From reversible quantum microdynamics to irreversible quantum transport, Physics Rep. 272 (1996), 1–59. ⟨124, 127⟩
[246] M. Reed and B. Simon, Methods of modern mathematical physics, Vol. 1, Functional Analysis, Academic Press, New York 1978. ⟨18, 47⟩
[247] W. Ritz, Recherches critiques sur l'électrodynamique générale, Annales Chemie et de Physique 13 (1908), 145–275. ⟨26⟩
[248] A. J. Roberts, Normal form transforms separate slow and fast modes in stochastic dynamical systems, Physica A 387 (2008), 12–38. ⟨198⟩
[249] H. P. Robertson, The uncertainty principle. Phys. Review 34 (1929), 163–4. ⟨34⟩
[250] L. Rosenfeld, Misunderstandings about the foundations of quantum theory, pp. 41–5 in: Observation and interpretation (S. Körner, ed.), Butterworths, London 1957. ⟨137⟩
[251] L. Rosenfeld, Letter to Belinfante, 24 July 1972, according to Footnote 243 in: S. Osnaghi, F. Freitas and O. Freire Jr, The origin of the Everettian heresy. Studies in History and Philosophy of Science B 40 (2009), 97–123. ⟨235⟩
[252] W. Rudin, Functional analysis. 2nd ed., McGraw-Hill, New York 1991. ⟨47, 48⟩
[253] D. Ruelle, On the asymptotic condition in quantum field theory, Helv. Phys. Acta 35 (1962), 147–63. ⟨122⟩
[254] B. Russell, The metaphysician's nightmare, in: Fact and fiction, Allen & Unwin, London 1961. ⟨234⟩
[255] S. Saitoh, Theory of reproducing kernels and its applications. Pitman Research Notes in Mathematics, Longman, 1988. ⟨90⟩
[256] S. Saitoh, Integral transforms, reproducing kernels and their applications. CRC Research Notes in Mathematics, Chapman & Hall, 1997. ⟨90⟩
[257] A. Salam, Molecular quantum electrodynamics in the Heisenberg picture: A field theoretic viewpoint, Int. Rev. Phys. Chem. 27 (2008), 405–48. ⟨208⟩
[258] L. L. Salcedo, Absence of classical and quantum mixing, Phys. Rev. A 54 (1996), 3657–60. ⟨160⟩
[259] W. Sandhas, Definition and existence of multichannel scattering states, Comm. Math. Phys. 3 (1966), 358–74. ⟨122⟩
[260] D. Sarason, Sub-Hardy Hilbert spaces in the unit disk, Wiley, New York 1994. ⟨83⟩
[261] G. Scharf, Finite quantum electrodynamics: The causal approach, 2nd ed., Springer, New York 1995. ⟨114, 119⟩
[262] M. Schirber, Focus: Windshield cracks hold secrets of impact, Physics 6 (2013), 48. http://physics.aps.org/articles/v6/48. ⟨204, 205⟩
[263] M. Schlichenmaier, Berezin–Toeplitz quantization for compact Kähler manifolds. A Review of Results, Adv. Math. Phys. 2010 (2010), 927280. ⟨102⟩
[264] M. Schlosshauer, Decoherence, the measurement problem, and interpretations of quantum mechanics, Rev. Mod. Phys. 76 (2005), 1267–305. ⟨124, 136, 174⟩
[265] M. Schlosshauer, Decoherence and the quantum-to-classical transition, Springer, New York 2007. ⟨VII, 124, 136, 174, 199, 234, 252⟩
[266] I. J. Schoenberg, Positive definite functions on spheres, Duke Math. J. 9 (1942), 96–108. ⟨66⟩
[267] I. J. Schoenberg, Linkages and distance geometry, I. Linkages, Indag. Math. 72 (1969), 43–52. ⟨64⟩
[268] B. Schroer, Infrateilchen in der Quantenfeldtheorie, Fortschritte der Physik 11 (1963), 1–31. ⟨237⟩
[269] E. Schrödinger, Der stetige Übergang von der Mikro- zur Makromechanik, Naturwissenschaften 14 (1926), 664–6. ⟨9, 71⟩
[270] E. Schrödinger, Might perhaps energy be a merely statistical concept? Il Nuovo Cimento

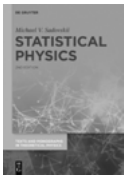
(1955–1965) 9 (1958), 162–70. ⟨137, 247⟩
[271] J. Schur, Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen, J. Reine Angew. Math. 140 (1911), 1–28. ⟨57⟩
[272] I. Schur, Über die Potenzreihen, die im Innern des Einheitskreises beschränkt sind, J. Reine Angew. Math. 147 (1917), 205–32. Fortsetzung, J. Reine Angew. Math. 148 (1918), 122–145. English translation pp. 31–59,61-88 in: I. Schur methods in operator theory and signal processing (I. Gohberg, ed.) Birkhäuser, Basel 1986. ⟨83⟩
[273] I. E. Segal, Quantization of nonlinear systems, J. Math. Phys. 1 (1960), 468–88. ⟨72⟩
[274] B. Simon, The classical limit of quantum partition functions, Comm. Math. Phys. 71(1980), 247–76. ⟨89⟩
[275] L. Sklar, Physics and chance, Cambridge Univ. Press, Cambridge 1993. ⟨31, 187⟩
[276] H. Spohn, Kinetic equations from Hamiltonian dynamics: Markovian limits, Rev. Mod. Phys. 52 (1980), 569–615. ⟨197⟩
[277] J. Stewart, Positive definite functions and generalizations, an historical survey, Rocky Mountain J. Math. 6 (1976), 409–34. ⟨70⟩
[278] M. L. Steyn-Ross and C. W. Gardiner, Quantum theory of excitonic optical bistability, Phys. Rev. A 27 (1983), 310–25. ⟨199⟩
[279] G. G. Stokes, On the composition and resolution of streams of polarized light from different sources, Trans. Cambridge Phil. Soc. 9 (1852), 399–416. Reprinted in [281]. ⟨148, 149⟩
[280] F. Strocchi, Complex coordinates and quantum mechanics. Rev. Mod. Phys. 38 (1966), 36–40. ⟨23⟩
[281] W. Swindell, Polarized light, Dowden 1975. ⟨264, 268⟩
[282] G. Szegö, Über orthogonale Polynome, die zu einer gegebenen Kurve der komplexen Ebene gehören, Math. Zeitschrift 9 (1921), 218–70. ⟨83⟩
[283] B. N. Taylor, Guide for the use of the international system of units (SI), NIST Special Publication 811, 1995 edition. http://physics.nist.gov/cuu/Units/introduction.html. ⟨31⟩
[284] B. N. Taylor and C. E. Kuyatt, Guidelines for evaluating and expressing the uncertainty of NIST measurement results, NIST Technical Note TN1297, 1994. https://www.nist.gov/pml/nist-technical-note-1297. ⟨183⟩
[285] G. Tenenbaum, Introduction to analytic and probabilistic number theory, Cambridge University Press, Cambridge 1995. ⟨42⟩
[286] W. E. Thirring, A course in mathematical physics III, Springer, New York 1987. ⟨18⟩
[287] H. Upmeier, Toeplitz operators and index theory in several complex variables. Birkhäuser, Basel 1996. ⟨90⟩
[288] N. Vandenberghe and E. Villermaux, Geometry and fragmentation of soft brittle impacted bodies, Soft Matter 9 (2013), 8162–76. ⟨205⟩
[289] J. Vignes, Review on stochastic approach to round-off error analysis and its applications, Math. Computers Simul. 30 (1988), 481–91. ⟨42⟩
[290] A. Vourdas, Coherent spaces, Boolean rings and quantum gates, Ann. Physics 373 (2016), 557–80. ⟨70⟩
[291] A. Vourdas, Coherent spaces, pp. 173–89 in: Coherent states and their applications, Springer Proc. Physics 205 (2018). ⟨70⟩
[292] D. Wallace, What is orthodox quantum mechanics? Unpublished Manuscript (2016). https://arxiv.org/abs/1604.05973. ⟨3⟩
[293] N. Wallach, The analytic continuation of the discrete series, I, II, Trans. Amer. Math. Soc. 251 (1979), 1–17, 19–37. ⟨68⟩
[294] S. Weinberg, Feynman rules for any spin, Phys. Rev. 133 (1964), B1318–22. ⟨118, 234⟩
[295] S. Weinberg, Feynman rules for any spin II, Phys. Rev. 134 (1964), B882–96. ⟨118⟩
[296] S. Weinberg, What is quantum field theory, and what did we think it is? Manuscript (1997).

arXiv:hep-th/9702027. ⟨210⟩
[297] S. Weinberg, Lectures on quantum mechanics, Cambridge Univ. Press, Cambridge 2013. ⟨241, 247⟩
[298] G. Wentzel, Zur Theorie des photoelektrischen Effekts, Zeitschrift f. Physik 40 (1926), 574–89. ⟨203⟩
[299] C. Wetterich, Non-equilibrium time evolution in quantum field theory, Phys. Rev. E 56 (1997), 2687–90. ⟨VII, 154⟩
[300] H. Weyl, Quantenmechanik und Gruppentheorie, Z. Phys. 46 (1927), 1–46. ⟨150, 220, 232, 254⟩
[301] J. A. Wheeler and W. H. Zurek (eds.), Quantum theory and measurement. Princeton Univ. Press, Princeton 1983. ⟨257, 258, 261⟩
[302] P. Whittle, Probability via expectation, 3rd ed., Springer, New York 1992. ⟨19, 31, 35, 45⟩
[303] A. Wightman, Hilbert's 6th problem, pp. 147–240 in: Mathematical developments arising from Hilbert problems (F. E. Brouwder, ed.), Proc. Symp. Pure Math. Vol. 28, Amer. Math. Soc., Providence, RI, 1976. ⟨137⟩
[304] E. P. Wigner, On unitary representations of the inhomogeneous Lorentz group, Ann. Math. 40 (1939), 149–204. ⟨209⟩
[305] E. Wigner, Die Messung quantenmechanischer Operatoren, Z. Phys. 133 (1952), 101–8. English translation: Measurement of quantum-mechanical operators, https://arxiv.org/abs/1012.4372 (See also Wigner's summary on p. 298 in: Wheeler & Zurek, Quantum theory and measurement, Princeton 1983.). ⟨237⟩
[306] Wikipedia, Born rule, Web document. https://en.wikipedia.org/wiki/Born_rule. ⟨179, 233⟩
[307] Wikipedia, Positive operator-valued measure (POVM). Web document. https://en.wikipedia.org/wiki/Positive_operator-valued_measure. ⟨236⟩
[308] Wikipedia, SIC-POVM, Web document. https://en.wikipedia.org/wiki/SIC-POVM. ⟨10⟩
[309] Wikipedia, Quantum state tomography. Web document. https://en.wikipedia.org/wiki/Quantum_state_tomography. ⟨125⟩
[310] B. G. Wybourne, Classical groups for physicists, Wiley, New York 1974. ⟨110⟩
[311] J. Yngvason, Localization and entanglement in relativistic quantum physics, pp. 325–48 in: The message of quantum science (P. Blanchard & Jürg Fröhlich, eds.), Springer, Berlin 2015. ⟨244, 245⟩
[312] S. Zaremba, Zarys Pierwszych Zasad Teoryi Liczb Calkowitych (in Polish). Akademia Umiejetności, Kraków 1907. ⟨70⟩
[313] G. Zauner, Quantendesigns – Grundzüge einer nichtkommutativen Designtheorie. Dissertation, Universität Wien, Österreich, 1999. http://www.gerhardzauner.at/documents/gz-quantendesigns.pdf English translation (including a new preface): Quantum Designs: Foundations of a non-commutative design theory, Ph.D. thesis, University of Vienna, Austria, 2010. http://www.gerhardzauner.at/documents/gz-quantumdesigns.pdf. ⟨10⟩
[314] H. D. Zeh, Physik ohne Realität: Tiefsinn oder Wahnsinn? Springer, Heidelberg 2012. ⟨5⟩
[315] W. M. Zhang and D. H. Feng, Quantum nonintegrability in finite systems, Phys. Reports 252 (1995), 1–100. ⟨108, 109, 194⟩
[316] W. M. Zhang, D. H. Feng and R. Gilmore, Coherent states: theory and some applications, Rev. Mod. Phys. 62 (1990), 867–927. ⟨68, 89, 102, 117⟩
[317] F. Ziegler, Answer to *Who coined the term "Born's rule"?*, History of Science and Mathematics, Q&A site, July 1, 2018. https://hsm.stackexchange.com/a/7488/3915. ⟨230, 258⟩
[318] W. H. Zurek, Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. 75 (2003), 715–75. ⟨174⟩

Authors

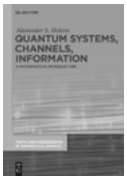
Index

Also of Interest



Statistical Physics
Michael V. Sadovskii, 2019
ISBN 978-3-11-064510-1, e-ISBN (PDF) 978-3-11-064848-5

Quantum Field Theory
Michael V. Sadovskii, 2019
ISBN 978-3-11-064515-6, e-ISBN (PDF) 978-3-11-064852-2



Quantum Systems, Channels, Information
A Mathematical Introduction
Alexander S. Holevo, 2019
ISBN 978-3-11-064224-7, e-ISBN (PDF) 978-3-11-064249-0

Elementary Particle Theory
Eugene Stefanovich, 2018

Volume 1: Quantum Mechanics
ISBN 978-3-11-049088-6, e-ISBN (PDF) 978-3-11-049213-2

Volume 2: Quantum Electrodynamics
ISBN 978-3-11-049089-3, e-ISBN (PDF) 978-3-11-049320-7

Volume 3: Relativistic Quantum Dynamics
ISBN 978-3-11-049090-9, e-ISBN: 978-3-11-049322-1